INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI

18877

I A R I 6
GIPC—S4—10 AR—21 6 49—1 000

PROCEEDINGS

OF THE

ROYAL SOCIETY OF LONDON

SERIES A—MATHEMATICAL AND PHYSICAL SCIENCES

VOL CLXII

LONDON
Printed and published for the Royal Society
By the Cambridge University Press
Fetter Lane E.C.4

15 October 1937

PRINTED IN GREAT BRITAIN BY
WALTER LEWIS, M.A.
AT THE CAMBRIDGE UNIVERSITY PRESS

CONTENTS

SERIES A VOL CLXII

No. A 908—1 September 1937

PAGE

Studies in Contact Potentials. The Condensation of Potassium and Sodium on Tungsten. By R. C. L. Bosworth and E. K. Rideal, F.R.S. (Plate 1) 1

Studies in Contact Potentials. The Evaporation of Sodium Films. By R. C. L. Bosworth. Communicated by E. K. Rideal, F.R.S. . . 32

The Electronic Specific Heat and X-Ray Absorption of Metals, and Some Other Properties Related to Electron Bands. By H. Jones and N. F. Mott, F.R.S. 49

The Elastic Stability of a Long and Slightly Bent Rectangular Plate under Uniform Shear. By D. M. A. Leggett. Communicated by R. V. Southwell, F.R.S. 62

The Absorption Spectra of the Monosulphides of Alkaline Earth Elements and their Latent Heats of Vaporization. By L. S. Mathur. Communicated by M. N. Saha, F.R.S. (Plates 2, 3) 83

The Angular Distribution of Photoelectrons from the K Shell. By F. F. H. Eggleston and L. H. Martin. Communicated by T. H. Laby, F.R.S. 95

The Absorption Spectrum of Formic Acid in the Vacuum Ultra-Violet. By W. C. Price and W. M. Evans. Communicated by R. G. W. Norrish, F.R.S. (Plate 4) 110

The Influence of the Difference of Orientation of Two Crystals on the Mechanical Effect of their Boundary. By Bruce Chalmers. Communicated by E. N. da C. Andrade, F.R.S. 120

Scattering of Slow Neutrons. By M. Goldhaber and G. H. Briggs. Communicated by Lord Rutherford, O.M., F.R.S. 127

No. A 909—15 September 1937

Quaternion Treatment of the Relativistic Wave Equation. By A. W. Conway, F.R.S. 145

Theory of Scattering of Protons by Protons. By Sir Arthur Eddington, F.R.S. 155

PAGE

Progressive Lightning. III—The Fine Structure of Return Lightning Strokes. By D. J. Malan and H. Collens. Communicated by Sir Charles Boys, F.R.S. (Plates 5–10) 175

Rotatory Dispersion in the Amine Series. I—*d-Spiro*heptane-Diamine and its Derivatives. By the late T. M. Lowry, F.R.S. and W. C. G. Baldwin. Communicated by R. G. W. Norrish, F.R.S. 204

Rotatory Dispersion in the Amine Series. II—*l*-Secondary Butylamine and its Derivatives. By W. C. G. Baldwin. Communicated by R. G. W. Norrish, F.R.S. 215

Rotatory Dispersion in the Amine Series. III—The Absorption Spectra of Diamines. By W. C. G. Baldwin. Communicated by R. G. W. Norrish, F.R.S. 228

Solution of Oseen's Equations for an Inclined Elliptic Cylinder in a Viscous Fluid. By D. Meksyn. Communicated by L. Bairstow, F.R.S. 232

Reversible Oxidation of Chlorophyll. By E. Rabinowitch and J. Weiss. Communicated by F. G. Donnan, F.R.S. 251

On the Nature of Atmospherics—V. By R. A. Watson Watt, the late J. F. Herd and F. E. Lutkin. Communicated by Sir George Simpson, F.R.S. 267

No. A 910—1 October 1937

The Behaviour of an Osglim Lamp. I—Introduction, Drift, Equilibrium and A.C. Bridge Measurements. By Lewis F. Richardson, F.R.S. (Plate 13) 293

The Behaviour of an Osglim Lamp. II—Oscillograph Methods, Steady Motion and Relaxational Oscillations. By Lewis F. Richardson, F.R.S. (Plates 14–16) 316

Stark Effect in Iron and the Contrast with Pole Effect. By S. F. Panter and J. S. Foster, F.R.S. (Plates 17, 18) 336

The Stark Effect in Hydrogen and Deuterium. By J. Stuart Foster, F.R.S. and Hawley Snell. (Plates 19, 20) 349

PAGE

The Influence of Hydrogen and Deuterium on the Thermal Decomposition of Diethyl Ether in the Low Pressure Region. By J. W. Mitchell and C. N. Hinshelwood, F.R.S. 357

Metallic Reflexion and the Surface Photoelectric Effect. By R. E. B. Makinson. Communicated by R. H. Fowler, F.R.S. 367

The Shape of the Continuous β-Spectrum of Thorium C.C′. By H. O. W. Richardson and Alice Leigh-Smith. Communicated by J. Chadwick, F.R.S. (Plates 21, 22) 391

Experiments on Iron Ammonium Alum in Alternating Fields at very Low Temperatures. By A. H. Cooke and R. A. Hull. Communicated by F. A. Lindemann, F.R.S. 404

The Absorption Spectrum of Liquid Bromine. By D. Porret. Communicated by F. G. Donnan, F.R.S. 414

Investigations of Infra-red Spectra—Absorption of some Hydroxy Compounds in the Region of $3\,\mu$. By J. J. Fox and A. E. Martin. Communicated by Sir Robert Robertson, F.R.S. 419

The Influence of Temperature on the Capture of Neutrons by Various Elements. By J. L. Michiels and J. A. Saxton. Communicated by G. P. Thomson, F.R.S. 441

The Thermoelastic Properties of Muscle and their Molecular Interpretation. By Kurt H. Meyer and L. E. R. Picken. Abstract . . . 450

No. A 911—15 October 1937

The Bakerian Lecture. Regularities and Irregularities in the Ionosphere—I. By E. V. Appleton, F.R.S. (Plates 23–28) 451

The Tests for Sampling Differences and Contingency. By Harold Jeffreys, F.R.S. 479

Researches on the Chemistry of Coal. VIII—Comparative Studies of the Macro-Constituents of Bituminous Coals. By William A. Bone, F.R.S. and Basil J. A. Bard 495

The Slow Combustion of Acetylene. By William A. Bone, F.R.S. and J. E. Carruthers 502

PAGE

The Diffusion of Pairs of Gases with Molecules of Equal Mass. By L. E. Boardman and N. E. Wild. Communicated by S. W. J. Smith, F.R.S. 511

Neutrons, Degeneracy and White Dwarfs. By D. S. Kothari. Communicated by M. N. Saha, F.R.S. 521

Nuclear Excitation and Disintegration Collisions involving Strong Interaction—I. By H. S. W. Massey and C. B. O. Mohr. Communicated by R. H. Fowler, F.R.S. 529

The Inhibition by Nitric Oxide of the Decomposition of Ethane. By L. A. K. Staveley. Communicated by C. N. Hinshelwood, F.R.S. 557

X-Ray Analysis of the Dibenzyl Series. IV—Detailed Structure of Stilbene. By J. Monteath Robertson and I. Woodward. Communicated by Sir William Bragg, O.M., P.R.S. 568

The Suspension of Solids in a Turbulent Stream. By E. G. Richardson. Communicated by W. E. Curtis, F.R.S. 583

Index 599

Studies in Contact Potentials

The Condensation of Potassium and Sodium on Tungsten

By R. C. L. Bosworth and E. K. Rideal, F.R.S.

(*Received* 18 *March* 1937)

[Plate 1]

Recent work on the adsorption of gases has emphasized both the importance and the difficulty of ensuring the cleanliness of metal surfaces. Methods which involve the "flashing" of a filament of the metal thus present several advantages over the more usual experimental arrangements. With such a filament several physical methods may be employed for investigating the effects of adsorbed films thereon, and a study of the accommodation coefficient (J. K. Roberts 1935) has already yielded much new information. Of the three electrical methods of examination of the film-covered surface, the thermionic work function, the photoelectric work function, and the contact potential, the last is to be preferred, for many films are mobile even at low temperatures, thus precluding use of the thermionic method. In others the photoelectric threshold is so far in the ultra-violet as to render its measurement difficult. From the contact potential the dipole moment of the chemi-adsorbed complex may be obtained. Its value and any variation with the packing density of the film-forming material are required both as a basis for evaluating the true form of the adsorption isotherm as well as for obtaining further insight into the mechanism of chemical reactivity at such surfaces.

The measurement of contact potentials by studying the infra-saturation currents in a thermionic tube originated in the work of Richardson and Robertson (1922). Oatley (1936) (cf. also Mönch 1928, 1930) obtained values for the contact potential difference of molybdenum, zinc and platinum against tungsten in quite satisfactory agreement with the best thermionic determinations of the work function; but, in general, the results obtained by contact-potential measurements are not always in satisfactory agreement with thermionic and photoelectric measurements, particularly when dealing with films. Although part of the discrepancy may be due, as

Compton and Langmuir (1930) have pointed out, to the fact that in inhomogeneous surfaces the contact-potential method measures the mean work function of the surface, while the other two methods measure the minimum work function, it would be well to survey the experimental technique to choose a procedure which will avoid the more usual difficulties.

Langmuir and Kingdon (1929) have laid down a set of conditions to be observed if reproducible results are to be obtained, but it would appear that still closer delimitation of the experimental conditions is required to render the method generally applicable. Let A and B be two filaments of any nature whatever, A hot and B cold; if A is kept at a fixed temperature the current to B is determined solely by the total potential difference between A and B. For negative (retarding) potentials the current, determined largely by the distribution in velocity of the emergent electrons, varies exponentially with the potential. At small positive potentials this form of variation changes over to the two-thirds power law required for a space-charge limited current, and finally at higher voltages to the Schottky relation, the whole curve current *v.* potential being perfectly definite and single valued. If, now, this current-volts characteristic is obtained with A at a certain temperature and B in a given state (α), and then the state of B is altered to (β) by forming or changing a film on the surface, and the current-volts characteristic is remeasured, this second curve will in general exhibit a lateral displacement relative to the first, and this displacement will measure the contact potential difference between B in the two states (β) and (α). If either or both of these states do not correspond to homogeneous films, the field irregularities due to the patches on B will average out at the surface A and the method will measure the mean contact potential of the inhomogeneous surface. If the film instead of being formed on B is formed on A, a similar but opposite displacement will be obtained, and from it the contact potential may be read off. If, however, the film formed on A is inhomogeneous the field at the surface of A—the emitting surface—will be irregular and the shape of the whole infra-saturation current-volts curve will be disturbed, so that the two curves obtained will no longer be parallel and no reliance may be placed on the contact-potential difference by the method above. In the work of Nelson (1931) filament A was a tungsten strip partially covered with barium and B a collecting anode of a special form. Nelson's current-volt characteristics with A in different states of activation exhibited non-parallelism, and the contact-potential differences could only be measured by the displacement in the points of breakdown of the exponential variation of current with retarding potential—a fact first established by Schottky (1914). The accuracy, however, with which the position of these kinks can be measured does not compare with the accuracy with which the displacement of two parallel curves can be measured. More recently, Anderson (1935) has carried out measurements forming barium films on the collector electrode, and his value is in much more satisfactory agreement with thermionic measurements than Nelson's.

It may therefore be stated that one further experimental condition for satisfactory contact potential measurements is that:

(1) The electron source should be an homogeneous surface and should remain constant during the experiments.

Another similar condition is:

(2) The electron source should be as nearly as possible equipotential.

This condition, i.e. using a short portion only of the emitting filament, is desirable, because if the potential difference between the source and the collector is spatially variable the relative contributions of the different parts of the source will change when the potential of the collector is changed, and change differently according as the potential change in the collector is homogeneous or patchy. Furthermore, it is desirable that the current should vary as rapidly as possible with the voltage, and the integral of the current-volts curve over a finite range will always vary less rapidly than the current itself.

One final condition is to be noted; it is only a modification of two of the conditions of Taylor and Langmuir (1933)—their conditions (2) and (3). In their work guard rings were used to eliminate end-effects, and they emphasized the care necessary to prevent photo-electrons from the metal plate vitiating the results. To minimize this danger it is desirable that:

(3) As little metal as possible should be used in the experimental tube, as every additional electrode is an additional source of both impurities and possible stray electrons.

The conditions above as well as Langmuir and Kingdon's requirements were all fulfilled in a recent paper by Reimann (1935), who measured the current-volts characteristics with the cold filament first clean and then covered with oxygen—no guard rings were necessary on account of the geometry of the arrangement, and his value of 1·70 V for the W-WO contact potential can be regarded as the first really satisfactory measurement of this quantity and the resolution of the difficulty of the impossibly large A found in Kingdon's (1924) experiments on the thermionic emission from a WO surface. Kingdon's value for the Richardson A and b were 5×10^{11} and 9·24 e-volts respectively, and as the theoretically maximum value for A is 120ρ (ρ the roughness factor) his experiments can only be reconciled with theory on the assumption that his b was falling with rising temperature, and Reimann was indeed successful in showing that the oxygen films were unstable over the temperature range used by Kingdon.

It is interesting at this stage to note that practically all the thermionic measurements on the electro-positive films agree in making A much less than 120; or, on the theory of the temperature variable work function of Becker and Brattain (1934) and Reimann (1934 *a*, *b*), in making the work function for these surfaces increase with increasing temperature. A large positive temperature coefficient of the work function may prove to be an intrinsic property of these films; it is, nevertheless, precisely the change to be expected if the positive films, like Kingdon's electronegative film of oxygen, were migrating away or evaporating at the temperature at which the thermionic measurements were obtained, and we are not without evidence from the photoelectric side that at least some of the electropositive films have a work function with a negative temperature coefficient (de Boer and Veenemans 1935; Suhrmann and Theissing 1932).

Another source of trouble arises since the current-volts curve of such a tube is strongly curved, making it a good rectifier for A.C. disturbances. At low voltages the curves are concave upwards, so that any A.C. will be rectified as a positive current, while at higher voltages the curves are concave downwards and A.C. is rectified as a negative current. The nett effect of A.C. on the cell is then to flatten out the current-volts curve and render the tube less sensitive to potential changes. Much of the trouble due to A.C. disturbances may be eliminated by the proper use of shielding.

A third point which arises is in connexion with the filament temperature. This, of course, should be kept constant, for the infra-saturation current is never, as sometimes indicated, independent of the temperature, although the temperature has progressively less effect at lower potentials. It is important to use as low a potential on the tube as will give a reasonable current. It seemed desirable before extending the method to the adsorption of gases to examine the behaviour of potassium and sodium films, as some information on their properties is already available from thermionic and photoelectric examination.

In a recent paper by one of us (Bosworth 1936*b*), it was shown that an electronegative layer between a film of potassium and a tungsten substrate increased the electropositive nature of the surface film, and it seemed desirable to investigate whether this result was due to an increase in the polarization of the surface atoms or in decreasing the depolarizing forces which have been shown to be active in concentrated films.

THE CONDENSATION OF POTASSIUM

To satisfy the conditions given above the two filaments are placed at right angles to one another and skew so that the centres clear by a distance of the order of a millimetre only, and then practically the whole of the current measured flows from a small length of a few millimetres only at the uniform centre of the hot filament to the centre of the cold filament.

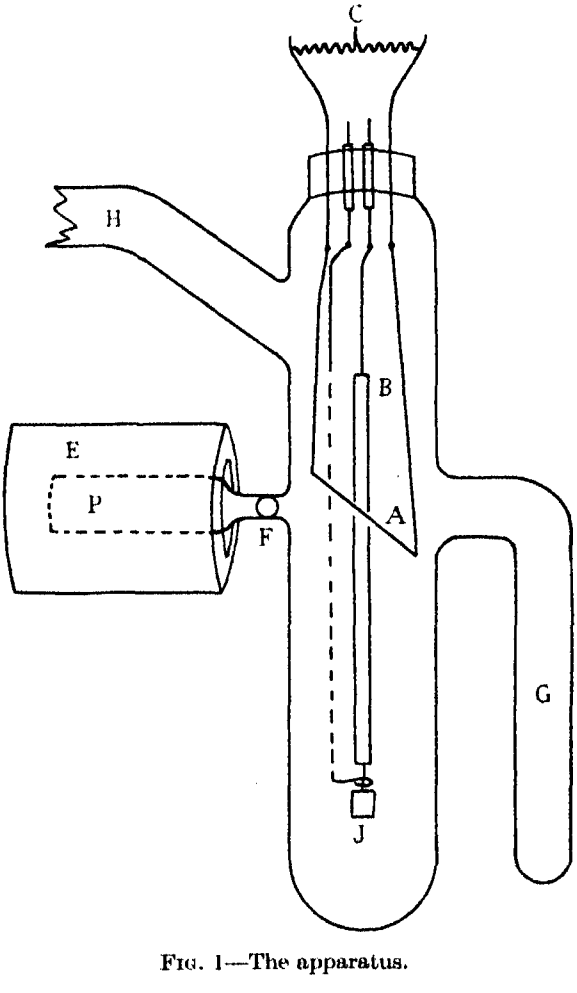


FIG. 1—The apparatus.

Incidentally this arrangement with suitable springs holding the filaments taut fulfils better than Reimann's the requirement latent in condition (5) of Langmuir and Kingdon—viz. that the clearance between the filaments should not change when they are heated. For the source A a horizontal filament of tungsten wire 0·10 mm. diam. was used, and for the collector B, a vertical filament of tungsten ribbon 2·02 mm. wide. Fig. 1 gives a sketch of the arrangement.

Filament B was held taut by a weight J and filament A by the natural elasticity in the leads. The external leads of A were connected by a centre tap resistance, the central point C provided an external point which had the same potential as the centre of the emitting surface, no matter what the heating current.

A cylindrical bulb P, sealed on by a narrow tube to the main bulb and placed in the same horizontal plane as the point of intersection of A and B, contained purified potassium. The vapour pressure in P could be raised to any desired value by means of a furnace E, and so a molecular beam of any desired intensity could be fired on to the centre of the strip filament. This beam could be cut off at will by means of a magnetically operated shutter F. That part of the beam which failed to deposit on the filament was trapped and condensed in the wide side tube G immersed in liquid air. The whole apparatus was shielded from electric disturbances by earthed metal screens and evacuated through the side tube H.

Method of Working

After the vessel had been evacuated, the potassium purified by multiple distillation and both filaments glowed out at as high a temperature as possible, the emission from A to B was measured over a range of potentials from -2 to $+6$ V, filament A being held at a fixed temperature. Filament B was then reflashed and hydrogen at a pressure of 10^{-2} mm. admitted to the vessel as soon as the flashing current was switched off. After 10–20 sec. the hydrogen was pumped out, and then 1 min. later A was flashed to remove all hydrogen from its surface and returned to the working temperature and the current to B measured over the same range of potentials as before. This process was then repeated admitting oxygen instead of hydrogen. The voltage displacements in the curves thus obtained gave the contact potentials of WH and WO surfaces against W. These two contact potentials have already been shown to be of the same sign and of the same order of magnitude by J. K. Roberts (1935), and the present work in which the changes in these two potentials is studied as the filament is gradually cleaned by heating at a high temperature to remove adsorbed and occluded gas has shown that the WH-W potential is always about 70% of the WO-W potential. As the clean-up of the surface proceeded both potentials increased at the same rate until after the filament had been heated to over 2400° K.

for such a length of time that the glass walls were warmed by the radiation to nearly 400° C., so that all the potassium and other impurities in the vessel were distilled into the liquid-air trap, the contact potential difference between the clean and the oxygenated or hydrogenated filament became constant and proved to be unaffected by any further heating of *A*. It was noted also that these potential differences were independent of the temperature of *A*. Fig. 2 shows the current-volts curve for the filament in the three states thus obtained, the filament *A* being held at a temperature of 2050° K. Curve *A* represents the emission with filament *B* bare, curve *B* with filament *B* covered with hydrogen, and curve *C*, *B* covered with oxygen. These curves give the following figures for the contact-potential differences:*

W-WH $-1{\cdot}26$ V,
W-WO $-1{\cdot}74$ V.

These figures may not represent the final values for these quantities and a more detailed determination is being undertaken, but they do, however, represent accurately the properties of the actual surfaces on which the deposition of potassium has been studied.

The method of operation consists in heating the oven around the potassium reservoir up to some standard temperature with the shutter closed, flashing filament *B*, and by means of curve *A* choosing some suitable potential such that the emission is small but not too small on switching on filament *A* at 2050° K. As soon as the current becomes steady it is noted and the shutter is opened; as the positive film builds up on *B* the current grows and curve *A* enables the current change to be interpreted as a change in the contact potential. It has been found, however, experimentally advisable, as the film builds up, to make the external potential connexions to *B* progressively more negative and so keep the current fairly small; in this way the readings obtained are far less susceptible to error arising from fluctuations in the temperature of filament *A*. The reading of the contact-potential difference after a given exposure to the potassium beam may be confirmed by closing the shutter and measuring a complete current-volts characteristic. In all cases where this has been done the curves obtained were parallel to the curves shown in fig. 2, so that one is given confidence in the readings obtained above by a single measurement of the current at a definite voltage.

The whole process is then repeated with the single difference that filament *B* is exposed to hydrogen immediately after flashing, so that now the change in contact potential is followed as the potassium film builds up on a hydrogen surface, curve *B* (fig. 2), of course, being used to convert the current change into a potential change. Finally, the formation of a potassium film over a WO surface is studied in a similar manner.

---

* The surface on the left is taken as negative to that on the right.

It is found that in all cases as the concentration of the film increases the contact potential rises, at first rapidly and then more slowly, finally becoming practically steady. This behaviour is not what one would at first

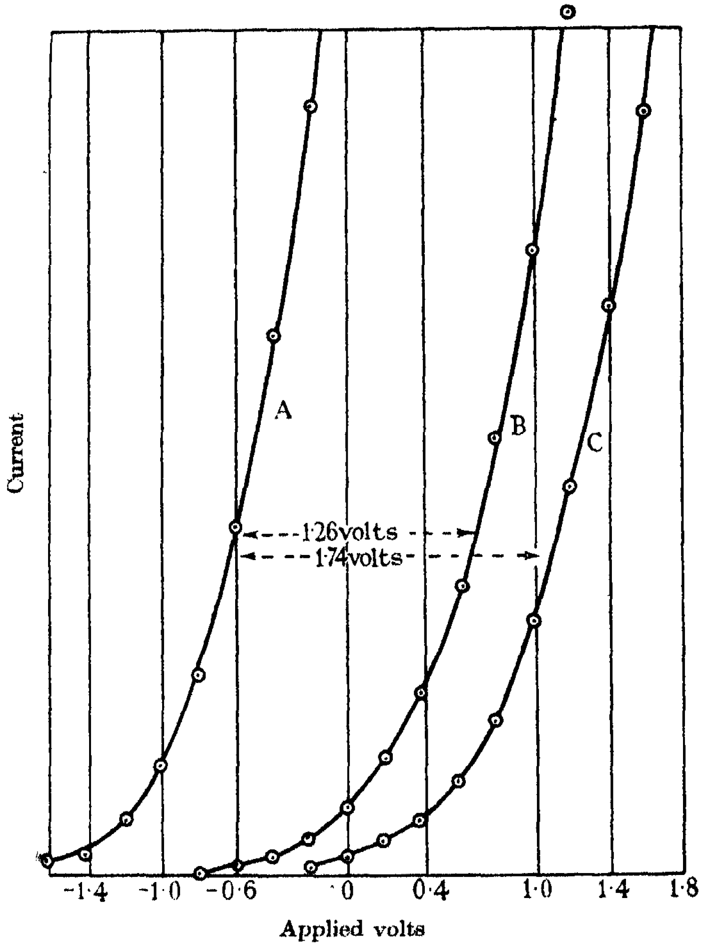


Fig. 2

sight expect, for the photoelectric threshold of different metals partially covered with potassium or other alkalis has been shown by Ives and Olpin (1929) to rise with increasing concentration, to reach a maximum, and then to fall asymptotically to a final value. But when it is recalled that, as shown recently (Bosworth 1936*a*), the surface-diffusion coefficient rises rapidly with concentration so that, at a certain concentration, the potassium might be expected to diffuse away from the central patch as fast as it is deposited. This, indeed, has proved to be the case. If, after depositing a not too concentrated film, the shutter is closed, the contact potential remains steady, showing that the central patch is stable and relatively immobile; but if a certain concentration is exceeded, then on closing the shutter the contact potential falls to a final steady value a few tenths of a volt lower than the contact potential at the moment of shutting off the beam. As far as can be ascertained by the somewhat crude method of shutting off the beam at different times, the concentration at which decay is first noticeable is much the same for WO, WH, or W surfaces. This observation lends support to the thesis advanced before by the authors—viz. that the rapid spreading observed in concentrated films is effected by the mutual repulsion of the adions.

The existence of this mobility sets a limit to the concentration at which the contact potential can be measured by the present method, and although the surface changes have been followed at three different rates of deposition (three different oven temperatures), it is doubtful whether with this apparatus the full monolayer value was ever reached.

Results

Tables I–III give the contact-potential difference of the surface against the same surface of zero concentration of potassium as a function of the time of exposure to the potassium beams arising from three different oven temperatures, the zero of time being measured in each case from the time of opening the shutter, and the contact potentials of K on WH films, for example, being measured from the WH surface taken as zero.

These are the final values, and exposure of the filament to the potassium beam for a further 10 min. produced no more change and no evidence for

TABLE I—OVEN TEMPERATURE 210° C.

Time (t) after opening shutter	Contact potential V_{12} on W V	on WH V	on WO V
0	0·00	0·00	0·00
30 sec.	0·10	0·10	0·10
1 min.	0·19	0·21	0·20
2 ,,	0·38	0·38	0·36
3 ,,	0·50	0·48	0·50
4 ,,	0·58	0·61	0·60
6 ,,	0·70	0·78	0·80
8 ,,	0·74	0·90	0·90

Table II—Oven Temperature 244° C.

	Contact potential V_{12}		
Time (t) after opening shutter	on W V	on WH V	on WO V
30 sec.	0·29	0·30	0·30
60 "	0·48	0·58	0·58
90 "	0·60	0·78	0·80
2 min.	0·72	0·94	1·02
3 "	0·90	1·22	1·36
5 "	1·16	1·52	1·80
7 "	1·32	1·78	2·12
10 "	1·50	2·10	2·54
15 "	1·80	2·36	2·76
20 "	1·84	2·26	2·86

Table III—Oven Temperature 287° C.

	Contact potential V_{12}		
Time (t) after opening shutter	on W V	on WH V	on WO V
30 sec.	0·82	1·06	1·26
60 "	1·10	1·70	2·16
90 "	1·66	2·12	2·54
2 min.	1·86	2·38	2·74
3 "	2·10	2·86	3·10
5 "	2·46	3·24	3·54
10 "	2·64	3·64	4·04
18 "	2·84	3·88	4·32
30 "	2·90	4·04	4·48

a maximum in the contact-potential curve was obtained, because, as stated above, surface migration prevents a layer thicker than a certain optimum thickness ever being built up. It would, of course, be easy to increase the intensity of the beam so as to cover all the available tungsten surface; but this, however, has been found to lead to the deposition of photosensitive films on the glass walls and filament leads and so give rise to parasitic currents. In short, no serious attempt has been made to study thick films, and the steady final value of the contact potential has been taken, provisionally, to be that due to the complete monolayer.

These contact potential differences are

W-WK	2·90 V,
WH-WHK	4·04 V,
WO-WOK	4·48 V.

But since for the electronegative films used

W-WH −1·26 V,
W-WO −1·74 V,

it follows that

W-WHK 2·78 V,
W-WOK 2·74 V.

Now the contact potential V_{12} between two surfaces (1) and (2) whose thermionic properties are expressed by the equations

$$I_1 = A_1 T^2 \exp(-V_1 e/kT)$$

and

$$I_2 = A_2 T^2 \exp(-V_2 e/kT),$$

can easily be shown to be

$$V_{12} = V_1 - V_2 + \frac{T}{11{,}600} \ln A_2/A_1,$$

so that the contact potential difference between two surfaces is equal to the difference in their electronic work functions plus a small Peltier term.

In the theory of the temperature variable work function which has been referred to above, A_1 and A_2 are to be taken as nearly equal, and the apparently large or small values found experimentally for certain surfaces are taken as due to a temperature variable factor in the corresponding V. Thus the Peltier terms above are quite small and may be neglected.

Figures for the thermionic properties of potassium on tungsten and oxygenated tungsten may be obtained from the results of Killian (1926). They are:

V_2 for WK 1·71 V,
V_2 for WOK 1·69 V.

And V_1, for tungsten at 300° K. on the assumption of an ideal value for A, is 4·56 V.

Thus the contact potentials are

W-WK 2·85 V,
W-WOK 2·87 V,

figures which agree satisfactorily with the measured values.

If, however, the experimental A's are accepted and the Peltier terms allowed for, the agreement between the calculated and measured contact-potential differences is not nearly so good. In the case of WHK films, the contact-potential method indicates a work function of 1·78 V which does not compare particularly well with the 1·34 V obtained photoelectrically, although the 1·66 V for the work function of WK compares very well with the photoelectric value of 1·72 (Bosworth 1936*b*).

It will be noted that for those concentrations for which the potassium film deposited is not mobile the value of θ, the surface concentration with respect to potassium, should be proportional to the time of exposure to the beam and to the beam intensity; assuming, of course, that the condensation coefficient for potassium behaves in the same way as for the caesium films studied by Taylor and Langmuir (1933), viz. remains unity for all values of θ. The relative intensities of the beams at the three different temperatures used are in the ratio of the vapour pressure of potassium at these temperatures, viz. 1 : 4·0 : 18·5, so that if the times (t) from Tables I–III be weighted in these ratios the three tables may be fitted into one; and from this a curve may be drawn showing how the contact-potential differences vary with θ from small values up to the point at which surface mobility at room temperature becomes appreciable, i.e. the point at which the contact potential at any given θ begins to vary with the rate of deposition—roughly the θ at which V_{12} for WK attains the value 1·8 V, for WHK 2·4 V and for WOK 2·8 V.

Analysis of Tables I–III shows that at very low concentration the addition of films of potassium of equal concentration to W, WH or WO surfaces produces equal changes in the contact potential, while at higher concentrations the addition of potassium to the surface produces an effect which decreases in magnitude as the electron affinity of the surface is decreased. The contact-potential difference produced by a film at concentration σ is related to the dipole moment μ of the adsorbed ions by the following equation:

$$V_{12} = 2\pi\mu\sigma.$$

Thus in very dilute films the dipole moment of the potassium adions is independent of the nature of the surface on which they are adsorbed, but, as the concentration increases, the dipole moment, under the influence of the mutually depolarizing forces acting between adions on the surface, falls. These depolarizing forces are most pronounced and begin to exert an influence at lowest concentration on the surface of lowest electron affinity, the W surface; and are least pronounced on the surface of highest electron affinity, the WO surface. As a result of this the dipole moments of the potassium ions on the surface, at all concentrations at which the depolarizing forces are operative, are in the order

$$\text{WOK} > \text{WHK} > \text{WK}.$$

In order to obtain an absolute measure of this dipole moment we require the absolute value of σ.

Practically every potassium atom (ionizing potential 4·32 V) striking a hot bare tungsten filament will be ionized, so that the strength of the ionic beam may be computed by heating filament *B* and observing the positive emission when the shutter is opened. The area of the patch formed on *B* may be estimated from the dimensions of the orifice of the oven, distance of the filament from the orifice and the distance and size of the visible patch of potassium ultimately formed in the side tube *G* (fig. 1): then, provided the patch formed is reasonably uniform, σ may be computed. In this way fig. 3 has been obtained. Fig. 3 shows the contact potentials against bare tungsten plotted against σ, from $\sigma = 0$ to $\sigma = 2 \times 10^{14}$. Higher values are not given, because the existence of surface mobility renders doubtful the measurement of V_{12} at higher values of σ.

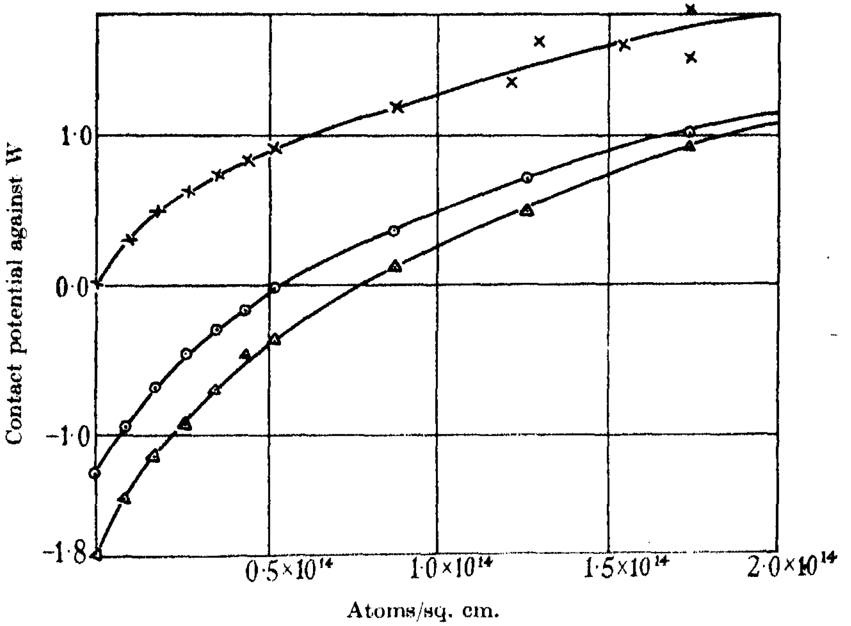


FIG. 3—Condensation of potassium. × K on W; ⊙ K on WH; ◬ K on WO.

From fig. 3 the values of σ may be calculated, and these values are given in Table IV on the assumption that the real values of σ are 1/1·3 of the apparent value, viz. that the roughness factor is 1·3 (Langmuir). There are addended also certain figures obtained by Bosworth for the dipole values of potassium films estimated from surface diffusion forces. These figures are given under W_D in Table IV.

TABLE IV

σ in 10^{14} atoms/sq. cm. of apparent surface	Dipole moment of K on WO	WH	W	W_D
0·03	22·0	22·0	22·0	—
0·06	22·0	22·0	22·0	23·0
0·10	21·5	21·5	21·0	—
0·12	21·2	21·2	20·7	21·0
0·18	20·7	20·6	18·6	—
0·24	20·2	19·7	16·6	—
0·36	18·7	18·0	13·7	—
0·48	17·8	16·1	12·2	13·0
0·60	16·6	14·8	11·0	—
1·0	13·5	11·8	8·7	—
1·5	11·2	9·8	7·4	—
2·0	9·8	8·3	6·2	—
2·4	—	—	—	6·0

These changes in dipole moment may be attributed to polarization induced by the electric field in the surface. Its measure is defined by $P = \mu_0 - \mu_\sigma$, where μ_0 is the dipole moment at infinite dilution.

The field exerted on any one pole of a selected dipole is the algebraic sum of the vertical components of the field due to the other dipoles, viz.

$$E = \mu \Sigma' \frac{1}{r^3}, \tag{1}$$

where r is the distance between any two dipoles on the surface. Different numerical evaluations of $\Sigma' \frac{1}{r^3}$ are obtained according to the surface structure postulated. If with Topping (1927) and Langmuir (1932) we assume a face-centred square lattice structure which expands uniformly on dilution we obtain

$$E = \frac{4\mu}{r^3} \Sigma_a \Sigma_b \frac{1}{(a^2+b^2)^{\frac{3}{2}}},$$

where a and b are the parameters determining the lattice co-ordinates

$$\infty \geqslant a \geqslant 1, \quad \infty \geqslant b \geqslant 0.$$

This equation reduces to $$E = 9{\cdot}02\mu\sigma^{\frac{3}{2}}, \tag{2}$$

since $$\sigma = \frac{1}{r^2}.$$

If, on the other hand, we regard the lattice points as fixed and the probability of any given point being filled simply as θ equation (1) reduces to

$$E = \frac{4\mu\theta}{r_0^3}\Sigma_a\Sigma_b\frac{1}{(a^2+b^2)^{\frac{3}{2}}} \quad \text{or} \quad E = 9{\cdot}02\mu\sigma\sigma_0^{\frac{1}{2}}, \tag{3}$$

where $\sigma_0 = \frac{1}{r_0^2}$, i.e. the surface concentration at $\theta = 1$.

In (2) all interaction between the adions and metal surface which tends to prevent the lattice structure from remaining uniform on expansion is ignored, whilst in (3) all interaction between different adions which tend to make the actual distribution more uniform than that assumed in calculating (3) is ignored. Thus the true field when both forces are operative must clearly lie between (2) and (3), and to a first approximation we may take the geometrical mean

$$E = 9{\cdot}02\mu\sigma^{\frac{5}{4}}\sigma_0^{\frac{1}{4}}, \tag{4}$$

until a more complete solution of the rather difficult summation of $\Sigma'\frac{1}{r^3}$ is obtained.

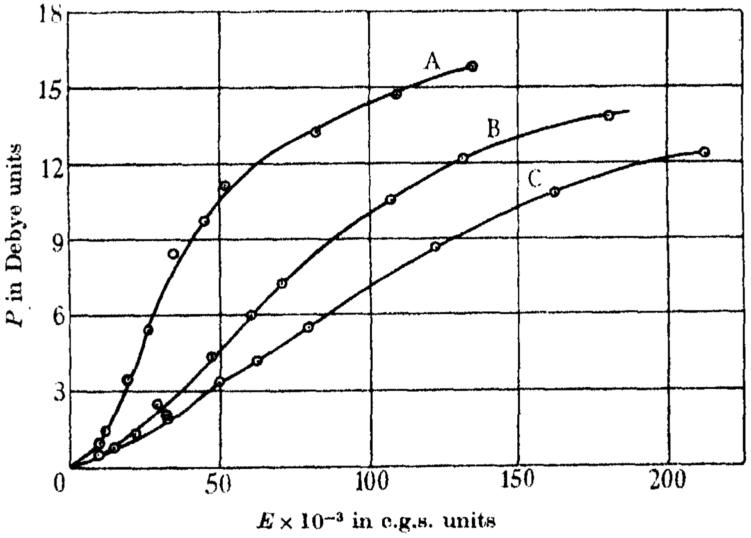


FIG. 4

In fig. 4 is shown $P = \mu_0 - \mu_\sigma$ plotted as a function of E given by (4) for films of potassium on tungsten A, tungsten hydride B, and oxide C respectively. The curves, as might be anticipated, are of the same form as the I, H curves obtained in the study of magnetization due to increasing field strength, and it is of interest to note that the Langevin equation is applicable to this system, viz.

$$P = P_\infty \coth \frac{2\alpha}{P_\infty} E - \frac{P_\infty}{2\alpha E},$$

where α is the polarizability of the films, and P_∞ the final value of P for large values of E must be less than μ_0, i.e. less than 22 Debye units.

In fig. 5 is shown a plot of P against $\coth \beta E - \frac{1}{\beta E}$, including all the experimental points on the three curves taking $P_\infty = 21 \times 10^{18}$ c.g.s. units, and β so chosen in each case as to give agreement at the point $\frac{P}{P_\infty} = 0{\cdot}5$.

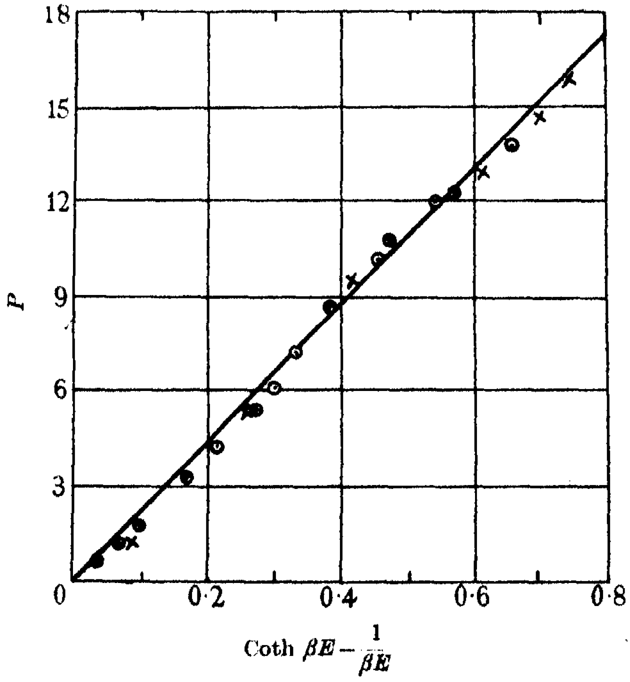
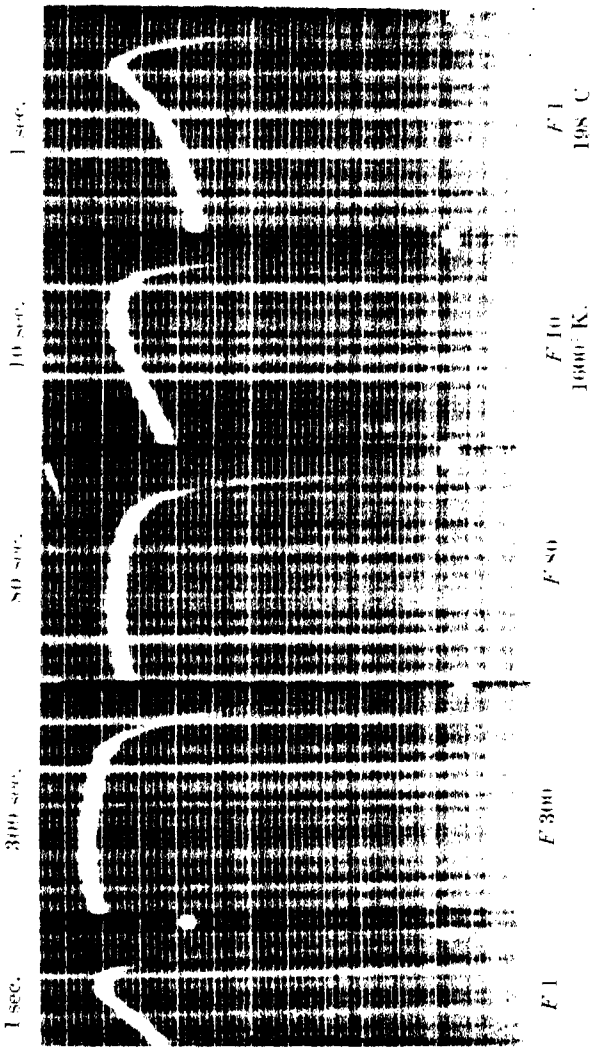


Fig. 5

The agreement is good, and the values of the polarizabilities obtained from the curve are

$\alpha = 300 \times 10^{-24}$ c.g.s. units for K on W,
$= 150$,, ,, for K on WH,
$= 100$,, ,, for K on WO,



(*Facing p.* 16)

figures which are considerably larger than $8 \cdot 1 \times 10^{-24}$ c.g.s. units given by Van Vleck (1932, p. 225) for gaseous potassium ions.

The magnitude of the depolarizing fields created by the film are worthy of note; thus for K on W at the highest concentration examined this field is 4×10^7 V/cm.

The effect of Hydrogen on the Potassium Films—The effect of exposing a clean tungsten filament to hydrogen is to lower its contact potential by 1·26 V. On the other hand, it has been shown (Bosworth 1936*b*) that there is no effect of exposure to molecular hydrogen on the photo emission of potassium. It was interesting, therefore, to examine the effect of exposing a tungsten filament partially covered with potassium to hydrogen. Table V shows the result of such a series of experiments. It gives the contact potential change δV_{12} produced by exposing to hydrogen a film of potassium on tungsten whose contact potential with respect to pure tungsten was originally V_{12}.

TABLE V

V_{12}	δV_{12}	$V_{12} - \frac{2 \cdot 90}{1 \cdot 26} \delta V_{12}$
0·00	− 1·26	2·90
1·00	− 0·60	2·40
2·00	− 0·40	2·92
2·78	− 0·04	2·88
2·90	0·00	2·90

The effect of hydrogen progressively decreases as the concentration of the potassium is increased, i.e. the hydrogen can only be chemiadsorbed on those lattice points which are bare. Likewise it is found that dilute films of potassium on WH or WO are unaffected by exposure to molecular hydrogen at low pressures because here there are no bare spots.

On the not unreasonable assumption that the dipole moment of four hydrogen atoms besides being of the opposite sign to a potassium adion is about 2/5 or more accurately 1·26/2·90 of the latter's value, the films above correspond to almost complete mixed film monolayers, for on reversing the sign of δV_{12} and multiplying by 2·90/1·26 and then adding to V_{12} the figures are obtained which are given in column 3 of Table V. These figures indicate that when four hydrogen ions are replaced by a potassium ion (four hydrogen ions occupy the same number of surface lattice points as one potassium ion), a fairly complete monolayer of potassium is formed.

We have, thus, evidence for the existence of a type of mixed film where electropositive and electronegative constituents occupy neighbouring lattice points in the same layer. The presence of such a film can only be shown by the contact-potential method which measures average work functions; for both the thermionic and photoelectric methods would measure the minimum work function of this surface and indicate only the presence of an unsaturated potassium film.

THE CONDENSATION OF SODIUM

Since the vapour pressure of sodium is much lower than that of potassium, a simplification of the apparatus was possible. The arrangement of cross-filaments previously described was sealed into a cylindrical glass bulb, and after evacuation and baking out, a little sodium was distilled *in vacuo* into the bottom of the bulb from a side tube which was then sealed off.

The experimental bulb could now be immersed in water or glycerine thermostats up to a temperature of 190° C. so as to adjust the sodium vapour pressure to any desired value. This pressure could also be made negligible by immersion of the bulb in liquid air. In some of the experimental tubes used, tungsten ribbon of the type employed before was used for the collector filament, while in others a wire filament 0·10 mm. diameter was used; a difference in the behaviour of these two types of surfaces will be alluded to below. When a filament is cleaned by flashing and exposed to sodium vapour at a constant pressure the thermionic current (I) which flows to that filament from a neighbouring source of constant intensity under a negative external potential grows at first rapidly and then more slowly, reaches a maximum, and then falls asymptotically to a final steady value. The curves (fig. 6) show five such successive studies of the rate of deposition. After each film had attained the properties indicated by the final point it was removed by a short flash at 1600° K. and the reading on the curves starting 1 min. to the right of the former curve taken. Fig. 6 shows that V_{12} increases less rapidly when σ is large than when σ is small, reaches a maximum at a certain σ, and thereafter decreases to a value independent of any further increase in σ. Contact-potential methods therefore like the photoelectric and thermionic methods (Ives and Olpin 1929; Taylor and Langmuir 1933; Becker 1929; Becker and Brattain 1933; and Nottingham 1932) indicate a maximum sensitivity at a certain optimum concentration.

Provided the sodium is removed by a short flash only the curves obtained in fig. 6 could be taken as typical either of the ribbon or the wire filament. When sodium is condensed on a ribbon filament however, somewhat different curves are obtained if the filament is heated to a high temperature prior to the deposition. The attached photographs (Plate 1) represent a string galvanometer record of the changes in I as the sodium film is formed on a surface which is cleaned by flashing at 1600° K. The curve on the extreme right shows the film building up after flashing for 1 sec., the successive curves after flashing for 10, 80 and 300 sec. respectively, and the last curve 1 sec. again, and satisfactorily reproduces the shape of the first. All the curves initially have the same shape—i.e. the rate of change of contact potential with time when the films are dilute does not depend on the time of flashing, but when the films are concentrated this rate does, indicating that the longer the filament is flashed the longer it takes to build up a film of optimum sensitivity on the surface, or more sodium is required to change the surface concentration by a given amount. Behaviour such as this has already been recorded (Bosworth 1935), and it was shown that films of sodium on tungsten thick enough to be photosensitive under white light tend to diffuse down crystalline cracks away from the surface. The diffusion rate increases both with the temperature and surface concentration, and at surface concentrations less than about 70% of the monolayer is quite inappreciable at room temperature. From this point of view it is possible to interpret these curves. The surface concentration increases regularly with the time until the work function of the surface begins to approach the minimum, and here the intergranular diffusion rate becomes appreciable, and the apparent rate of change of surface concentration is the difference between the rate of arrival of sodium at the surface from the vapour phase

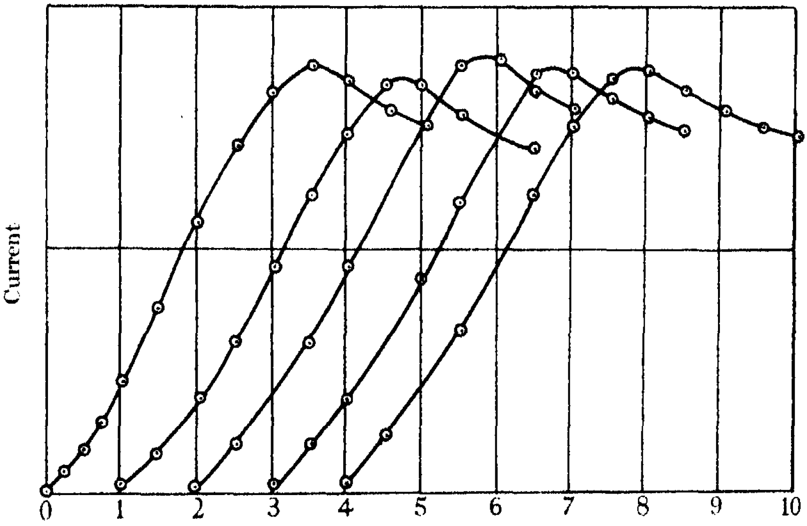


FIG. 6

and its rate of removal by diffusion. This last rate is determined by the inner concentration of sodium which is, in turn, determined by the time of flashing. Heating to 1600° K. evaporates all the surface sodium practically instantaneously, but the sodium from the inner structure only as fast as it can diffuse to the surface. Prolonged heating therefore at a high temperature leaves the inner structure far emptier than a shorter heating, so that in the former case it will apparently take more sodium to reach any given surface concentration than in the latter case. The time taken to reach the maximum in the V_{12} curve should therefore be a measure of the amount of evaporation from the inner surface and should therefore increase both with time of flashing as actually indicated in the photographs and the temperature of flashing. This second point is indicated in fig. 7, which shows *I-t* curves of the same type as those in the photographs (Plate 1). Here, however, a much lower vapour pressure of sodium has been used, and the results were obtained by direct reading. The essential parts of the condensation curves alone are shown, i.e. the immediate vicinity of the maximum. Curves in fig. 7*a* show the condensation on a surface prepared by

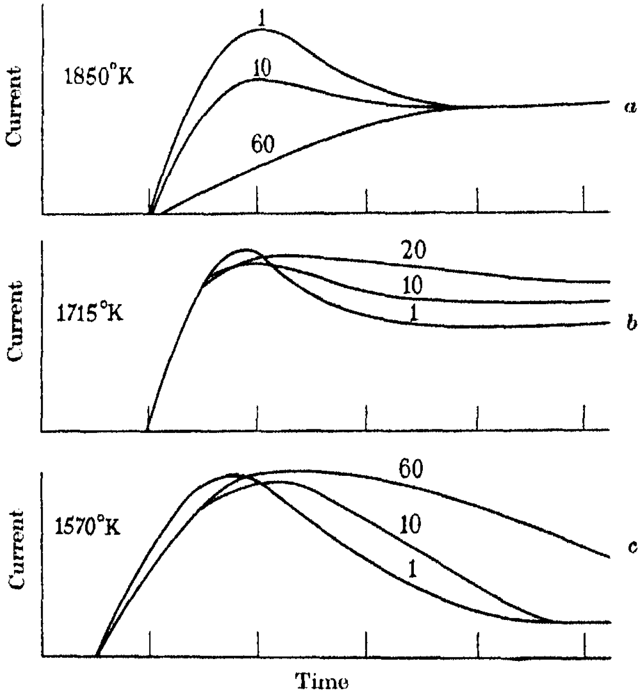


Fig. 7

flashing for 1, 10 and 60 sec. respectively at 1850° K., curves in fig. 7*b* for 1, 10 and 20 sec. respectively at 1715° K., and curves in fig. 7*c* for 1, 10 and 60 sec. at 1570° K., and the more pronounced effect of heating at higher temperatures is clearly indicated. It would be difficult to draw any accurate quantitative information from these curves as the phenomena are clearly complex. A comparison in the relative shift in the position of the maxima, however, indicates that the evaporation rates for the three temperatures given are roughly in the ratio of 15 : 10 : 6, i.e. the activation energy for the process determining the speed of this evaporation is thus of the order of 0·8 e-volt. The slowest process involved in this evaporation will clearly be the migration of the sodium to the surface, for the actual evaporation from the surface itself is known to be very fast at far lower temperatures, and the figure given above must represent the activation energy for migration over an almost bare surface. It is to be noted that this figure is of the same order of magnitude as the 0·7 e-volt given for the activation energy for the migration of potassium on a bare surface (Bosworth 1936*a*).

The Vapour Pressure of Sodium

The fact that the contact potential of the surface will only remain constant if there are no adsorbable molecules striking the surface can be used to measure the rate at which molecules accumulate on the surface, i.e. the pressure. As a means of measuring pressure this method has all the sensitivity of Langmuir and Kingdon's (1925) positive-ion method and is of more general availability. The former method, which consists in measuring the positive-ion current flowing from a hot wire in the alkali vapour, will measure a vapour pressure of 10^{-11} mm. easily; but the method is, however, practically restricted to substances with an ionization potential less than 4½ V and of such a stability that they are not decomposed by a wire at 1000° K. In practice this means restriction to potassium, rubidium, and caesium. The measurement of the drift in contact potential, however, is available for all substances capable of being adsorbed on any wire that can be cleaned, and no restriction whatever is placed on the ionization potential of stability. The method, however, demands a knowledge of the condensation coefficient and the number of molecules per unit area in a film of any given properties—information which is not always available.

The behaviour of sodium on tungsten is probably similar to the deposition of caesium. Taylor and Langmuir (1933) showed the condensation coefficient to be unity. There is still uncertainty as to the actual number of molecules per unit area in the film of minimum work function. This has often been taken as corresponding to the monolayer; but in the case of caesium both the experiments of Becker (1926) and those of Taylor and Langmuir (1933), when proper allowance is made for the roughness factor, indicate that a minimum work function is attained when θ is about 70%, and de Boer (1935) has calculated that for barium films maximum thermionic emission is attained at $\theta = 0{\cdot}29$, so that we may anticipate for the sodium films that the maximum in the V_{12}-θ curve will be attained at some value of θ less than unity, and the absolute values of the vapour pressures are best determined by fitting the vapour pressure-temperature curve to that obtained at higher temperatures and pressures by the effusion method (Edmondson and Egerton 1926). The effusion and contact-potential method have this point in common, that they effectively count the number of atoms evaporating from the surface and so are affected in the same way by the occurrence of association in the gaseous state. A Na_2 molecule in the gaseous state will give the same contribution to the pressure as a Na atom, but will be counted by the method herein used as equivalent to two Na atoms. Consequently the contact-potential method measures the pressure which the vapour would exert were the vapour completely dissassociated, and if an appreciable number of Na_2 molecules are present this would be sensibly higher than the true pressure.

The rate at which atoms strike a surface is given by the Knudsen effusion equation:

$$r = 3{\cdot}54 \times 10^{22} \frac{P}{\sqrt{(MT)}} \text{ atoms/sq. cm./sec.}; \tag{1}$$

P is the pressure in mm. of mercury, T the absolute temperature, and M the atomic weight.

Let τ be the time in seconds taken to reach a surface concentration σ.

Then
$$\begin{aligned} 1/\tau &= r/\sigma \\ &= 3{\cdot}54 \times 10^{22} \frac{P}{\sqrt{(MT)}\sigma}. \end{aligned}$$

If we assume that the sodium film has the same lattice structure as a caesium film, σ corresponding to a monolayer (allowing for a roughness factor of 1·3 in a well-aged tungsten wire) is $4{\cdot}8 \times 10^{14}$, so that we may rewrite the equation above by substituting for σ, $4{\cdot}8 \times 10^{14}\theta$, or

$$\frac{1}{\tau} = 7{\cdot}35 \times 10^{7} \frac{P}{\sqrt{(MT)}\theta},$$

or taking $M = 23$,

$$\frac{1}{\tau} = 1{\cdot}56 \times 10^{7} \frac{P}{\theta\sqrt{T}}. \tag{2}$$

The experimental procedure is now to surround the bulb with a bath at a constant temperature, and after waiting about half an hour for the vapour to attain its equilibrium value to give the collector filament a short flash to clean its surface and then to note the rate of change of contact potential. At higher temperatures the complete curve is obtained and the time taken to reach the maximum noted. At the lower temperatures where this time is of the order of hours it is necessary only to observe part of the V_{12}-τ curve and compare that part of the curve with the corresponding parts of the higher temperature curves. Table VI expresses the results of a series of such determinations, in which the times τ are those required to attain the maximum contact potential at T° K.

TABLE VI

T° K.	τ (sec.)
273	180,000
285	57,000
299	24,000
312	3,000
322·5	1,200
336	330
356·5	30·0
373	7·5
386	2·4
406	0·57
434	0·110
443	0·056

The points are plotted as $\log \tau$ against $1000/T$ in fig. 8, and it will be noted that for liquid sodium the best experimental plot is represented by the equation

$$\log \tau = \frac{5148}{T} - 12{\cdot}89, \tag{3}$$

and for solid sodium

$$\log \tau = \frac{5333}{T} - 13{\cdot}38. \tag{4}$$

These lines intersect at $\tau = 370^\circ$ K. which is therefore the melting point of the sodium (I.C.T. 370·5° K.), and attests to the purity of the sample used. At 373° K. the vapour pressure of sodium, according to the equation given by Edmondson and Egerton (1926), is

$$1{\cdot}25 \times 10^{-7} \text{ mm.}$$

Using equation (2) in the form

$$\theta = \frac{1{\cdot}56 \times 10^7 \tau P}{\sqrt{T}},$$

it is readily seen that θ corresponds to 0·75, or the film gives it maximum contact potential when it is three-quarters saturated.

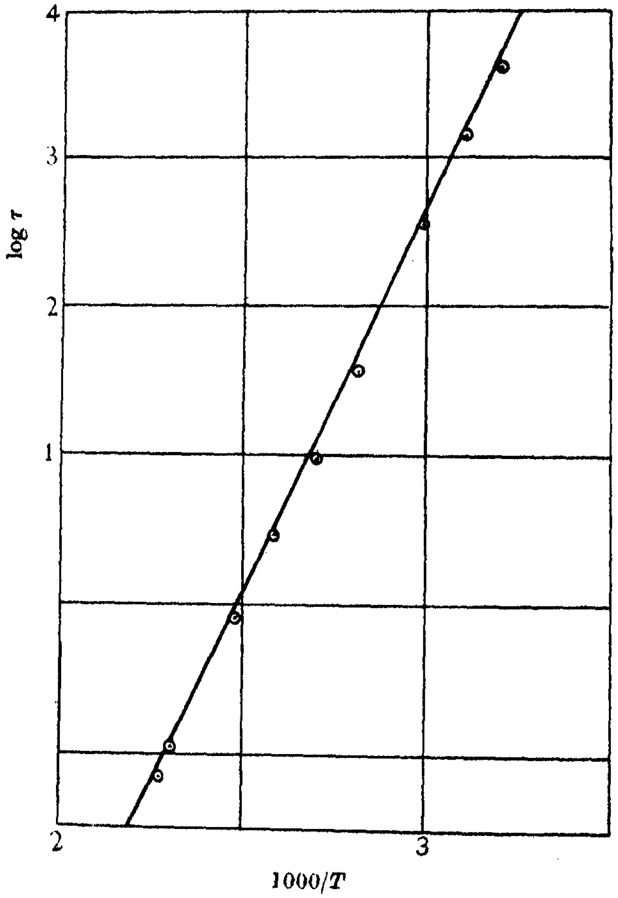


FIG. 8

Using this fact equations (3) and (4) may be expressed with the pressures in mm. of mercury. These equations are

Above 370° K.:

$$\log P = 7{\cdot}26 - \frac{5248}{T}. \tag{5}$$

Below 370° K.:

$$\log P = 7{\cdot}68 - \frac{5410}{T}. \tag{6}$$

The vapour-pressure equation for the liquid phase may be compared with that given in the "International Critical Tables" for the mean of several observations over a higher temperature range, viz.

$$\log P = 7{\cdot}553 - \frac{5392}{T},$$

and the agreement must be regarded as satisfactory.

The difference in the terms involving T in equations (5) and (6) is $\frac{1}{4{\cdot}575}$ of the difference in heats of evaporation from solid and liquid sodium at the temperature of 370° K.; viz. the latent heat of fusion. From the equations this is

$$4{\cdot}575 \times 162 = 740 \text{ cal.}$$

The latent heat of fusion is known to be 600 cal. by more direct measurements, and so the method for measuring vapour pressures which has been developed above is, when due precautions are taken, capable of giving accurate measurements.

The roughness factor was assumed for the sake of these experiments to be the same as that recorded by Taylor and Langmuir, viz. 1·3, for a wire which had been similarly treated. When condensation on tungsten ribbon was studied it was necessary to assume a far higher specific surface or a roughness factor of the order of 3–4 in order to obtain the same absolute value of the vapour pressure. This roughness factor tended to decrease on ageing.

The Absolute Measure of the Contact Potential as a Function of the Concentration

For the calibration of current reading against contact potential three current-volts curves were used, and they are shown in fig. 9. Curve *A* represents the emission to an anode which has been exposed to sodium vapour (saturated at 373° K.) so long that its work function has become constant. Curve *B* is the emission obtained when the bulb is immersed in liquid air and all the sodium removed by heating the filament to a high temperature, and curve *C* is the emission obtained on admitting a little O_2 by heating a filament coated with BaO_2 in a side tube. The voltage displacement *BC* should represent the W-WO potential, and this from fig. 4 is

1·82 V. This value must be regarded as the upper limit of this potential for the reason that curve *B* represents not pure tungsten but tungsten with a possibility of a very dilute film of sodium covering it. In the same way the figure given earlier in this paper, 1·74 V, represents the potential difference between WO and W covered with a dilute residual film of oxygen and is therefore a lower limit to the W-WO potential. The voltage displacement between curves *A* and *B* is 2·48 V, which is therefore the contact-potential difference between the sodium film in equilibrium with the vapour at 373° K. and tungsten. The film of maximum work function is 0·26 more positive than this or the $WNa_{(max)}$-W potential is 2·74 V. If we refer potentials to the WO surface as standard, taking the mean of the upper (1·82 V) and lower limit (1·74 V) of the W-WO potential, the true W potential is 0·04 more negative than that shown in fig. 4, and the best measures for the contact potential differences measured above become

$WNa_{(sat)}$-W 2·52 V,
$WNa_{(max)}$-W 2·78 V,
and also $WO_2Na_{(max)}$-W 2·82 V.

This last measurement was taken from the maximum potential attained when the oxidized filament was exposed to sodium vapour.

We can also examine how the contact potential changes with θ, and fig. 10 shows V plotted as a function of the surface concentration from the former I versus t curves with the help of the calibration implicit in fig. 9, and the fact that V is a maximum when σ (apparent) is $3{\cdot}6 \times 10^{14}$.

From fig. 10 the dipole moment of the sodium adions on the surface may be calculated with the help of the equation

$$V = 2\pi\mu\sigma \text{ e.s.u.'s},$$

or

$$\mu \text{ (the dipole moment)} = \frac{V}{2\pi\sigma} \times 10^{18} \text{ Debye units.}$$

The results of such calculations are shown in Table VII, where μ is given as a function of σ.

TABLE VII

σ in 10^{14} atoms/sq. cm.	μ in Debyes	σ in 10^{14} atoms/sq. cm.	μ in Debyes
0·04	22·6	1·1	9·0
0·09	22·0	1·5	7·8
0·22	19·0	1·9	6·0
0·40	15·8	2·7	5·4
0·60	12·7	3·6 (monolayer)	3·8
0·75	11·0		

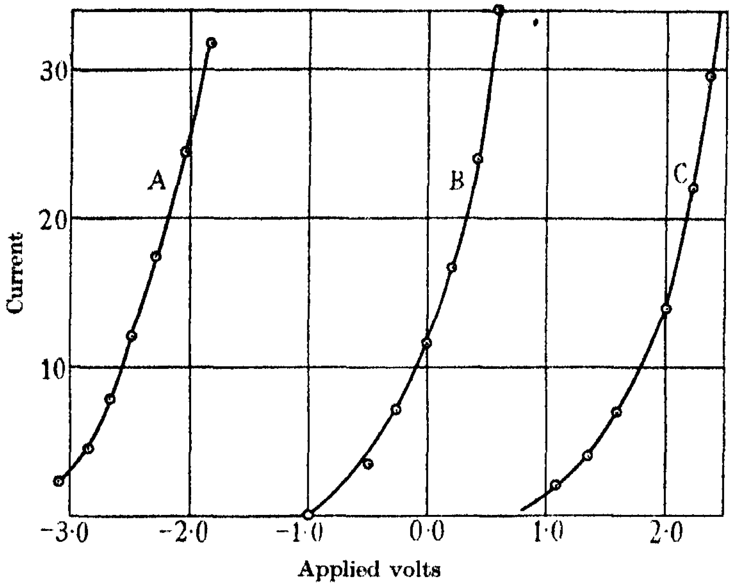

30
20
10
Current
A
B
C
−3·0
−2·0
−1·0
0·0
1·0
2·0
Applied volts

FIG. 9

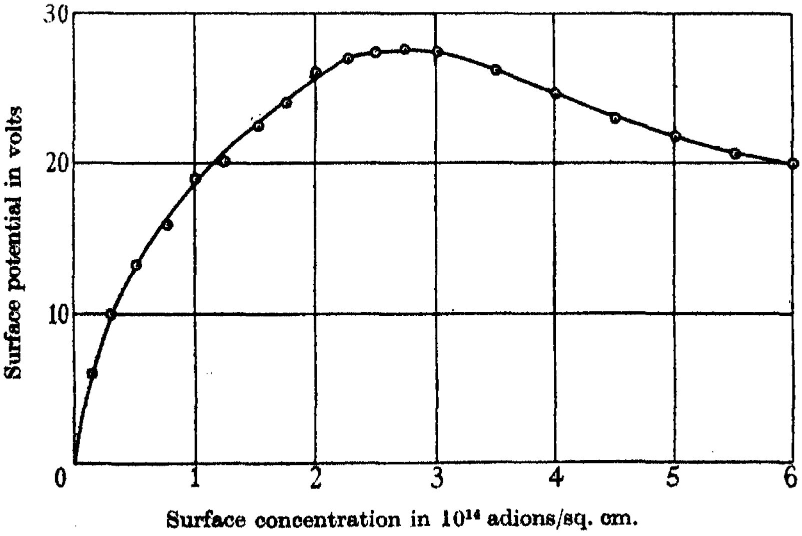

30
20
10
0
Surface potential in volts
1
2
3
4
5
6
Surface concentration in 10^{14} adions/sq. cm.

FIG. 10

The surface concentration has been calculated in atoms per unit area of apparent surface, so that $\sigma = 2{\cdot}7 \times 10^{14}$ corresponds to the surface of minimum work function and $3{\cdot}6 \times 10^{14}$ to the monolayer. It will be seen from the table that μ decreases with increasing σ, and at the monolayer is only one-sixth of its value in very dilute films. This variation of σ with μ is a measure of the polarizability. Fig. 11 shows P, the polarization in Debye units, plotted against E, the field in 10^7 V/cm. The curve obtained is of the same general form as those recorded before, but it is evident, however, that the polarizability of sodium is far less than that of potassium. The α corresponding to fig. 11 is only 11×10^{-23} c.g.s. units, or about one-third the value recorded for potassium films. This particular fact is all the more remarkable when it is recalled that the dipole moment attributed to the undepolarized sodium adions is within experimental error the same as that of the potassium adions, and the work functions of the final surface differ only by about 1/5 V. The polarizability is a far more specific property of the surface film than any other property so far examined.

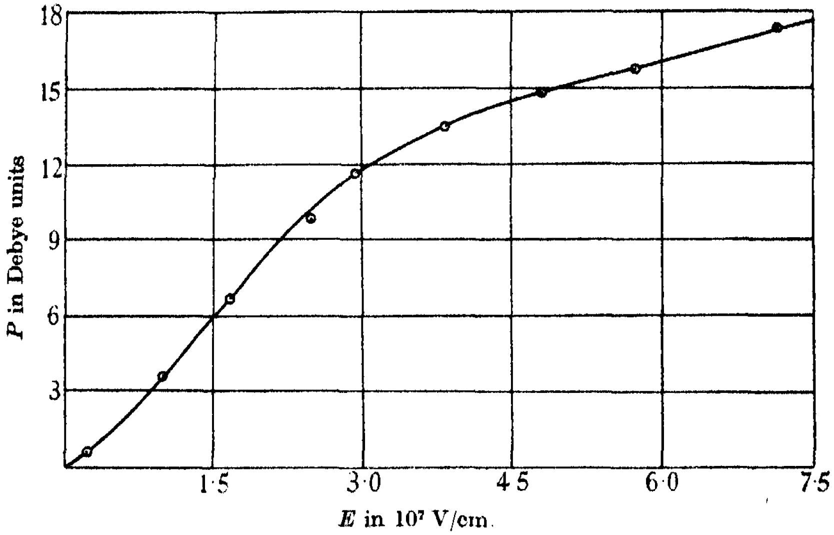


Fig. 11

Some Properties of Sodium on Oxide Films

Reference to this has already been made in giving the value of the WO_2Na-W potential which was the maximum potential attained when a filament oxidized to a blue colour by heating to 900° K. for half a minute in oxygen at a few millimetres pressure was exposed to sodium vapour. The properties of films on such oxidized tungsten ribbon have been investigated by Powell and Mercer and by Frank, and they have drawn attention in particular to the very large capacity of these oxide films for such metals as caesium, thallium and indium.

In the present experiments the oxidized filament was exposed to sodium vapour saturated at 100° C., and the rate of change of surface potential recorded. The results are given in Table VIII.

TABLE VIII

Surface potential V	Time in min.	Surface potential V	Time in min.
0·00	0	2·92	40
0·13	1	3·82	80
0·22	2	4·33	120
0·60	5	4·43	200
1·08	10	4·54	300
1·67	20	4·60	400
2·40	30	4·50	600

When the time of 400 min. to reach the maximum is compared with the 7·5 sec. required when the film was deposited on micro-crystalline tungsten the extremely porous nature of the oxide may be appreciated.

One of the authors (R. C. L. B.) is greatly indebted to the Royal Commissioners for the award of a Senior Exhibition of 1851.

SUMMARY

Experimental conditions to be observed when the method of contact potentials is used to examine the properties of films on metal surfaces are given, films of sodium and potassium at different concentrations are examined, and from the contact-potential differences the dipole moments of the adions have been calculated. In all cases studied the dipole moment decreases with increasing film concentration. In the case of sodium for example in dilute films the electric moment is 22·6 Debye units. This value falls with rising θ, for the surface of minimum work function is 5·4 Debye units, and for the monolayer only 3·8 Debye units.

This decrease of the dipole moment with increasing θ is attributed to the existence of depolarizing fields produced by the mutual interaction of the adions. These fields are shown to be of the order of 10^7 V/cm. and of the right

sign to induce a dipole of the opposite sense to the natural dipole. The relationship between induced dipole and field is worked out and the polarizability of the adions computed.

Potassium films are studied on W, WH and WO surfaces and the dipole moment of the potassium adion proves to be the same on all three surfaces when the films are dilute. With increasing concentration however, the dipole moments decrease most rapidly on the W surface and least rapidly on the WO surface, thus the polarizabilities of the K ion on the various surfaces are in the order

$$\mathrm{W} > \mathrm{WH} > \mathrm{WO}.$$

The disturbing influence produced by surface migration at high concentrations makes it impossible, however, to study the dipole changes above an apparent film concentration of $2{\cdot}0 \times 10^{14}$ atoms/sq. cm.

Figures are given for what is believed to be an almost complete film of potassium over the bases W, WH and WO and are compared with the thermionic measurements of Killian and the photoelectric measurements of Bosworth. For the WK film, excellent agreement is obtained, for the composite films not so good an agreement.

A study is also made of the adsorption of hydrogen on tungsten surfaces partially covered with potassium. It is found that a complete covering of hydrogen, potassium, oxygen, or a mixed film excludes any further adsorption of hydrogen, and that on partially covered surfaces hydrogen may be adsorbed only on those lattice points which are bare and by such adsorption a composite surface is formed which is a complete film.

When a tungsten filament is cleaned and exposed to sodium at a fixed pressure, the work function falls to a minimum and then rises slowly to final steady value. If the condensation be, however, studied on a ribbon it is found that the specific surface is far greater than the wire and that the number of sodium atoms required to change the surface concentration from $\theta = 0{\cdot}8$ to $\theta = 1{\cdot}0$ is increased by prolonged heating at a high temperature. The results are interpreted in terms of intergranular diffusion. The effect is even more pronounced in oxide films of the type investigated by Powell and Mercer (1935), where it is shown that the effective area is 3000 times the apparent surface.

From the rate of change of surface potential when the filament is exposed to sodium vapour the pressure of that vapour may be determined. The vapour pressure of solid sodium over the range 285–370° K. is found to be best represented by

$$\log p = 7{\cdot}68 - 5410/T,$$

and for liquid sodium over the temperature range 370–443° K.

$$\log p = 7{\cdot}26 - 5248/T.$$

The contact potentials of the films of minimum work function ($\theta = 0{\cdot}75$) against clean tungsten are

$$\begin{array}{ll} \text{W-WNa} & 2{\cdot}78\ \text{V}, \\ \text{W-WO}_2\text{Na} & 2{\cdot}82\ \text{V}. \end{array}$$

REFERENCES

Anderson, P. A. 1935 *Phys. Rev.* **47**, 958–64.
Becker, J. A. 1926 *Phys. Rev.* **28**, 341–61.
— 1929 *Trans. Amer. Electrochem. Soc.* **55**, 153–75.
Becker, J. A. and Brattain, W. H. 1933 *Phys. Rev.* **43**, 428–50.
— — 1934 *Phys. Rev.* **45**, 694–705.
Bosworth, R. C. L. 1935 *Proc. Roy. Soc.* A, **150**, 58–76.
— 1936*a* *Proc. Roy. Soc.* A, **154**, 112–23.
— 1936*b* *Trans. Faraday Soc.* **32**, 1369–75.
Compton, K. T. and Langmuir, I. 1930 *Rev. Mod. Phys.* **2**, 123–242.
de Boer, J. H. 1935 "Electron Emission and Adsorption Phenomena", p. 104.
de Boer, J. H. and Veenemans, C. F. 1935 *Physica, 'sGrav.*, **2**, 529–34.
Edmondson, W. and Egerton, A. 1926 *Proc. Roy. Soc.* A, **113**, 520–33; 533–42.
Ives, H. E. and Olpin, A. R. 1929 *Phys. Rev.* **34**, 117–28.
Killian, T. J. 1926 *Phys. Rev.* **27**, 578–87.
Kingdon, K. H. 1924 *Phys. Rev.* **24**, 510–15.
Langmuir, I. 1932 *J. Amer. Chem. Soc.* **54**, 2798–832.
Langmuir, I. and Kingdon, K. H. 1925 *Proc. Roy. Soc.* A, **107**, 61–79.
— — 1929 *Phys. Rev.* **34**, 129–35.
Mönch, G. 1928 *Z. Phys.* **47**, 7–8, 522–41.
— 1930 *Z. Phys.* **65**, 3–4, 233–43.
Nelson, H. 1931 *Physics*, **1**, 84–93.
Nottingham, W. B. 1932 *Phys. Rev.* **41**, 793–812.
Oatley, C. W. 1936 *Proc. Roy. Soc.* A, **155**, 218–34.
Powell, C. F. and Mercer, R. L. 1935 *Philos. Trans.* A, **235**, 101–24.
Reimann, A. L. 1934*a* *Phys. Rev.* **45**, 898.
— 1934*b* *Nature, Lond.*, **133**, 833.
— 1935 *Phil. Mag.* **20**, 594–607.
Richardson, O. W. and Robertson, F. S. 1922 *Phil. Mag.* **43**, 162–74.
Roberts, J. K. 1935 *Proc. Roy. Soc.* A, **152**, 445–63.
Schottky, W. 1914 *Ann. Phys., Lpz.*, **44**, 1011–32.
Suhrmann, R. and Theissing, H. 1932 *Z. Phys.* **73**, 11–12, 709–26.
Taylor, J. B. and Langmuir, I. 1933 *Phys. Rev.* **44**, 423–58.
Topping, J. 1927 *Proc. Roy. Soc.* A, **114**, 67–72.
Van Vleck, J. H. 1932 "Theory of Electric and Magnetic Susceptibilities", Oxford: Clarendon Press.

Studies in Contact Potentials
The Evaporation of Sodium Films

By R. C. L. Bosworth

(*Communicated by Eric K. Rideal, F.R.S.—Received* 18 *March* 1937)

In this paper the contact potential method of Bosworth and Rideal has been applied to the study of films of sodium on tungsten at elevated temperatures. Since the filament on which the sodium film under observation was deposited was heated most conveniently by the passage of an electric current, it was important to ensure that no error was introduced in the measurement of the contact potential due to the potential drop along the filament caused by the heating current. This difficulty has been overcome by two methods, first by employing a rotary split commutator permitting alternate measurements of the contact potential and passage of the heating current, a device suitable for relatively slow changes in film properties, but unsuitable for rapid changes necessitating the employment of a string galvanometer. In the second method the system was so arranged that the electron-collecting portion of the filament was effectively at its geometrical centre, and the geometry of the filament system was unaffected by changes in the filament temperatures. Under these conditions a Wheatstone's Bridge arrangement enabled the potential difference down the filament to be eliminated and the correct position of the collecting area ensured by noting the position of no change in the recorded contact potential on reversal of the heating current through the filament.

We may note that in reversible changes the surface potential changes on heating but returns immediately on cooling to its original value; such changes are assumed therefore to be unaccompanied by any change in the surface concentration (θ) and are intrinsic thermal properties of the film at a constant θ. The surface potential of the tungsten filament itself changes with varying temperature, and any irregularities due to maladjustment of the apparatus can be detected in this way.

The Variation of the Surface Potential of Sodium Films with Temperature

If the equation for the thermionic emission from a surface (i) is represented by the Richardson equation

$$I = A_i T^2 e^{-h_i/T},$$

where h_t is the Becker and Brattain (1934) heat function and A_t the intercept, then it may be shown by thermodynamical reasoning due first to Richardson (1916) that the contact potential difference (V_{12}) between two surfaces (1) and (2) is

$$V_{12} = k/\epsilon(h_1 - h_2) + \frac{kT}{\epsilon} \ln A_1/A_2 \tag{1}$$

$$= \frac{k}{\epsilon}(b_1 - b_2), \tag{2}$$

where b_1, b_2 are the work functions (temperature variable) of the two surfaces, defined by

$$I = 120 T^2 e^{-b_t/T}.$$

Now if dV is the potential difference between two surfaces of the same chemical nature differing by temperature dT, it follows from equation (2) that

$$\frac{dV}{dT} = -\frac{k}{\epsilon}\frac{db}{dT} = -\frac{1}{5040}\log\frac{120}{A}. \tag{3}$$

In the case of the pure tungsten surface the current-volts characteristic with the filament at 1710° K. was first obtained, the filament cooled to 80° K. and the characteristic remeasured; finally this was repeated at 1230° K. The contact potentials obtained against W at 80° K. were:

1230° K.	$-0{\cdot}085$ V,
1710° K.	$-0{\cdot}115$ V,

from which

$$\frac{dV}{dT} = -7{\cdot}0 \times 10^{-5},$$

and $A = 54$ in good agreement with the value of 60 found in the best thermionic determinations of this quantity. Raising the temperature of the collector still higher results, however, in a more rapid change in the apparent contact potential as attested by the figures below:

Temperature °K.	Contact p.d. against cold W V
1775	$-0{\cdot}12$
1975	$-0{\cdot}36$
2100	$-0{\cdot}64$

This particular effect has been discussed by Kuhn (1932) and is to be attributed to the negative emission from the collector filament building up a space charge around the filament and so lowering its apparent potential. The incidence of this effect puts an upper limit to the temperature at which contact potentials may be studied. This upper limit may be raised somewhat by raising the temperature of the source and thus making the emission of the collector negligible in comparison with the source. In general, however, it has been found in all the studies at present undertaken that 1900° K. is about the upper limit to the temperature at which it is possible to study the surface potential of pure tungsten with correspondingly lower temperatures for surfaces with electro-positive contamination. Bearing these points in mind, the study of the temperature effects on the surface potential of sodium films was undertaken.

A sodium film was formed over the filament by distilling a little sodium into the tube containing the filaments. Excess sodium was frozen out by immersion of the bulb in liquid air and the current-volts characteristic re-measured. It was 2·74 V positive. The filament was next heated to 600° K.; the surface potential at first rose to a higher positive value (dV/dT is *therefore positive*) and fell slowly to a lower value as evaporation proceeded. When the surface potential had become reasonably steady it was remeasured (+ 2·58 V) and the filament cooled to 80° K., when the surface potential fell sharply to + 2·42 V. The temperature of the filament was then raised till a little more sodium had evaporated and the surface potential measured at a number of lower temperatures. This process was repeated till all the sodium had evaporated, and the points obtained are represented in the curves of fig. 1, which show the surface potential plotted against temperature at a number of fixed values of the surface concentration (the curve for $\sigma = 0{\cdot}05 \times 10^{14}$ has been raised 1 V to fit it on to the same diagram as the others, viz. for $T = 600°$K., $V = 0{\cdot}32$ V not 1·32 as shown).

The higher temperature points (between 800 and 1000° K.) were not obtained by the above method, for at these temperatures the evaporation rate of sodium films at all concentrations is quite rapid. The points were obtained from string galvanometer records of the evaporation processes. All these records showed an initial kick to a more positive potential before the steady drift towards negative potentials sets it, and by extrapolating back the evaporation curves to the time of switching on the heating current it is possible to make an estimate of the value which the surface potential would possess at the temperature of evaporation were the surface adions prevented from evaporating. The higher temperature points are then necessarily somewhat inaccurate, and since the filament takes a finite time to reach its final temperature and because the evaporation process is, as will be seen later, most rapid in the initial stages, the error will be one of defect. The true thermal changes at high temperatures are probably larger than indicated by fig. 1. We are therefore justified in saying both that the temperature coefficient of the work function is negative (a hot surface is positive relative to a cold) and that this temperature coefficient itself increases rapidly with temperature. Thus if we adopt the picture provided by equation (3) the values of *A* for these surfaces are all greater than 120, so that the results here recorded are quite the antithesis of the thermionically determined values of *A* for films of caesium, rubidium and potassium

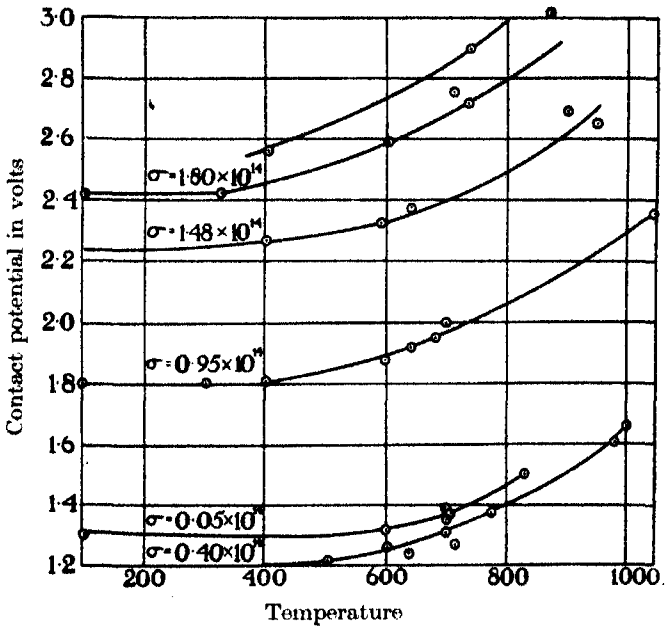


FIG. 1

(Langmuir and Kingdon 1925; Killian 1926), where figures smaller than 120 were always recorded.

At all temperatures, however, at which thermionic emission from sodium films is measurable, the vapour pressure in equilibrium with the film is also measurable, so that any thermionic measurements on sodium films must necessarily be carried out on a film whose θ is decreasing with rising temperature and will consequently be expected to give a (spurious) value of *A* probably much less than 120. That the true value of *A* is greater than 120 is indicated by the photoelectric work of de Boer and Veenemans (1935), where they showed that films of sodium on tungsten below a certain concentration were more sensitive to white light at higher temperatures.

Thus the work function must decrease with rising temperature and A be greater than 120. In explanation of their results, de Boer and Veenemans assume that on account of thermal agitation the adions are at the higher temperatures, somewhat further on the average from the surface, so that the dipoles which they, in conjunction with their electric images, constitute are of somewhat larger moment.

Instead of assuming a random distribution of adions above the surface we may postulate a series of higher surface energy levels such as suggested by Lennard-Jones (1932) to account for the mobility of surface films and picture the adions as distributed between the normal surface energy level and a higher level more distant from the surface, and the activation of an adion to this level would accordingly be accompanied by an increase in its dipole moment and a further lowering of the work function.

On this view the variation of μ, the dipole moment, with temperature will be

$$\mu = \mu_0 + \delta\mu e^{-b/T},$$

where $\delta\mu$ is the increase in the dipole moment effected when an adion is raised to the excited level and b is the activation energy for this jump.

The form of variation of V with temperature is then of the type

$$V = 2\pi\sigma(\mu_0 + \delta\mu e^{-b/T}),$$

or if δV is the change in surface potential produced by the activation of some of the adions,

$$\delta V = 2\pi\sigma\delta\mu e^{-b/T}. \qquad (4)$$

This equation fits the curves of fig. 1 with tolerable accuracy, and b is found, as expected, to increase when σ is decreased. Thus for

$$\sigma = 1{\cdot}8 \times 10^{14}, \quad b = 0{\cdot}22\,\text{e-volt},$$

and for

$$\sigma = 0{\cdot}95 \times 10^{14}, \quad b = 0{\cdot}38\,\text{e-volt}.$$

These figures are somewhat lower than the activation energy recorded (Bosworth 1935) for the migration of sodium, but when it is recalled that all the tendency was for errors in defect in the higher temperature measurements the agreement may be regarded as satisfactory. There is yet another source of error in the high temperature region—the effect of the thermionic emission from the collecting filament. This will begin to make the filament apparently too negative at temperatures of the order of

$$\frac{1900(4{\cdot}6 - V)}{4{\cdot}6}.$$

$\simeq$ 900° K. for the more concentrated films and provides yet another reason for believing that the b's recorded above are too small. Indeed, the general form of the curves shown in fig. 1 could be explained if the filament was emitting positive ions comparable with or greater than the electronic emission and so building up a positive sheath around the filament. We know, however, that this is not true (the positive ion emission from concentrated films is negligible at all temperatures) and have assumed above that these ions are building up a sheath around the filament while remaining essentially adions, by the process of being activated to an excited level; and this level we are at the present moment justified in identifying with the mobile level (0·3 e-volt from the normal level in concentrated films).

Accordingly at 600° K. it is to be expected that a fraction

$$\exp(-1400/600) = 0{\cdot}1$$

of the surface ions will be in the higher level. This fraction increases the contact potential by 0·16 V, and consequently from equation (4) their dipole moment must have been increased by

$$\frac{5{\cdot}30 \times 0{\cdot}16}{0{\cdot}1 \times 1{\cdot}8} \text{ Debye units}$$
$$= 4{\cdot}7 \text{ Debye units.}$$

In other words the arm of the electric moment has been increased by 1 A, equivalent to a movement of the ion 0·5 A away from the surface.

THE EVAPORATION OF SODIUM FROM TUNGSTEN

The evaporation has been studied by raising the filament to a known temperature and then following the negative drift (after the initial positive change) of the surface potential. Curves such as those shown in fig. 2 are obtained showing the contact potential change as a function of the time and temperature. These curves represent pure evaporation only when the experimental tube has been cooled in liquid air. In the presence of sodium vapour the evaporation rapidly comes to an end at a surface concentration at which the evaporation rate is equal to the condensation rate. Such curves as those shown in fig. 2 may with the help of the relation between θ and V obtained before (Bosworth and Rideal 1937) be converted into θ-time curves. Since by far the greatest change in θ occurs in the first few seconds, it is preferable to plot them as $d\theta/dt$, θ curves, which possess the advantage of enabling experimental results obtained in the presence of sodium vapour to be used by adding to the apparent $(-d\theta/dt)$ observed from evaporation curves the $+d\theta/dt$ (with the sign reversed) observed from the condensation curve obtained at the same vapour pressure.

The contact-potential method in its present form is restricted to the use of films more dilute than the film giving the maximum contact potential, i.e. to θ's below 0·75, although with the aid of a three-filament tube at present under consideration it is hoped to extend this range.

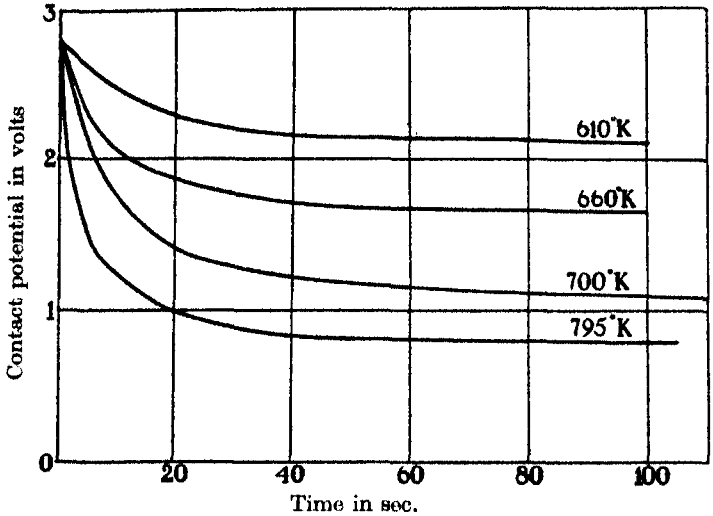


FIG. 2

It is convenient to measure surface concentrations in terms of the "f" units introduced by Becker (1932). By f is to be understood

$$f = \frac{\theta}{\theta_m} \text{ or } \frac{\sigma}{\sigma_m},$$

where θ_m and σ_m are measures of the surface concentrations of films of the minimum work function. $f = 1$ then signifies the most concentrated film the present method is capable of studying.

The curves shown in fig. 3 show f as a function of $\frac{1}{f}\frac{df}{dt}$ for evaporation at various temperatures. The results shown were obtained from string galvanometer records of the evaporation process. For temperatures below about 900° K. the evaporation is slow enough to be followed by hand, and by this means another series of similar curves has been obtained down to temperatures of 610° K. These low-temperature results will be shown later to be consistent with the higher temperature ones, but could not be conveniently shown in fig. 3.

The curves in fig. 3 are of immediate significance in that df/dt is a measure of the pressure in equilibrium with a film of concentration f at a fixed temperature T. In point of fact it may easily be shown that

$$\sigma_m \frac{df}{dt} = 3{\cdot}54 \times 10^{22} p (MT)^{-\frac{1}{2}} \tag{5}$$

or

$$\frac{df}{dt} = 2{\cdot}06 \times 10^{7} p T^{-\frac{1}{2}}.$$

The heat of evaporation of the film may be computed as

$$R \frac{d \log p}{d \frac{1}{T}}$$

$$= R \frac{d \ln (df/dt \, T^{\frac{1}{2}})}{d \frac{1}{T}}, \tag{6}$$

and fig. 4 shows $\log \left(\frac{1}{f} \frac{df}{dt} T^{\frac{1}{2}} \right)$ plotted against $10{,}000/T$ as required by equation (6). From the graph it is clear that at any fixed f the values of p in equilibrium with film at a temperature T are represented by equations of the type

$$\log p = a - \frac{b}{T},$$

where a and b are both functions of f. From the lines in fig. 4 the values of a and b given in Table I were obtained; $\log p$ is of course

$$\bar{7}{\cdot}314 + \log \frac{df}{dt} T^{\frac{1}{2}}$$

$$= \bar{7}{\cdot}314 + \log \frac{1}{f} \frac{df}{dt} T^{\frac{1}{2}} + \log f.$$

TABLE I

f	A	b	Q cal./mol.
0·0	7·40	7000	32,000
0·2	7·00	6250	28,500
0·4	6·92	5900	27,000
0·6	6·40	5000	23,000
0·8	6·10	4400	20,000
1·0	5·80	3750	17,000

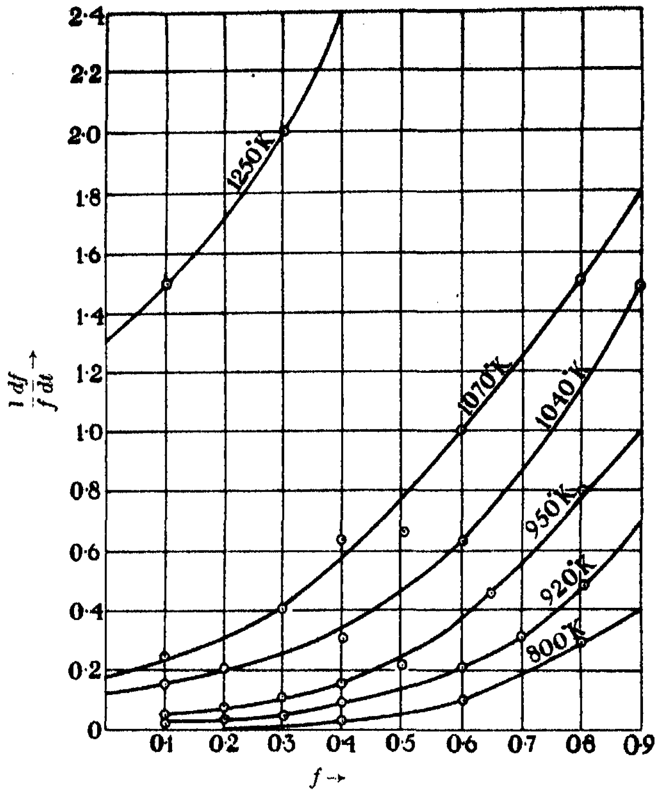

2·4
2·2
2·0
1·8
1·6
1·4
1·2
1·0
0·8
0·6
0·4
0·2
0
$\frac{1}{f}\frac{df}{dt}$ →
1250°K
1070°K
1040°K
950°K
920°K
800°K
0·1
0·2
0·3
0·4
0·5
0·6
0·7
0·8
0·9
f →

Fig. 3

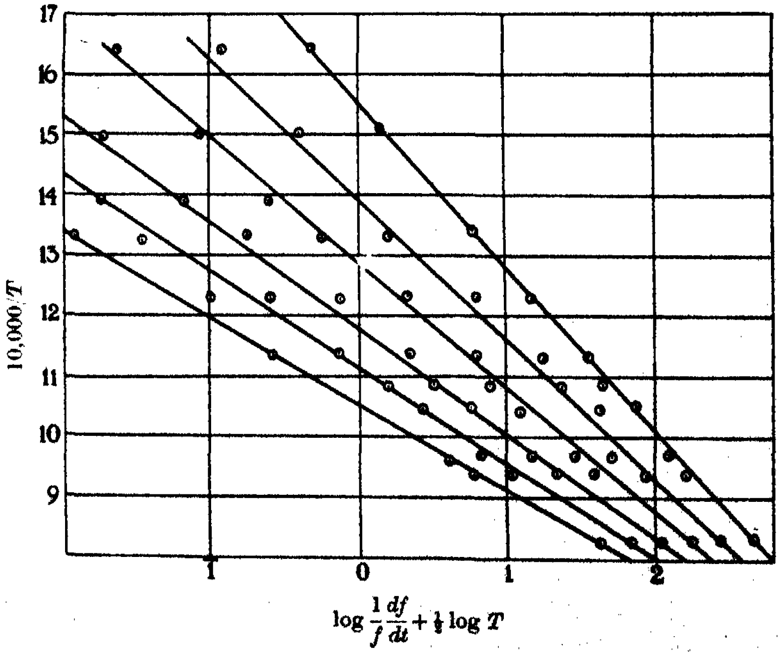

17
16
15
14
13
12
11
10
9
10,000/T
1
0
1
2
$\log \frac{1}{f}\frac{df}{dt} + \frac{1}{2}\log T$

Fig. 4

The lines in fig. 4 may be expressed in the form

$$\log \frac{1}{f}\frac{df}{dt}T^{\frac{1}{2}} = A - \frac{b}{T} = A - \frac{Q}{2 \cdot 3RT},$$

where the values of A, b and $Q = 2 \cdot 3Rb$ are given in Table I.

We note that the heat of evaporation for sodium from a very dilute film is 32,000 cal. or 1·4 e-volts, which figure steadily decreases as the surface concentration increases, becoming only 17,000 cal. or 0·75 e-volt for the film of minimum work function. This last figure is considerably less than the 25,000 cal. found for the heat of evaporation of solid sodium (Bosworth and Rideal 1937), and consequently at still higher film concentrations $f > 1$ this heat must increase with surface concentration.

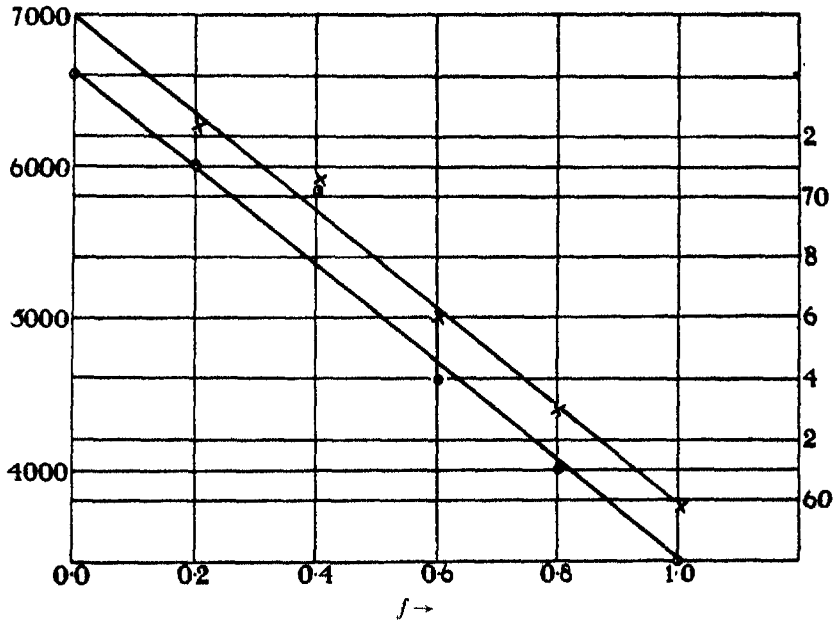


Fig. 5

The variations of A and b with f indicated in Table I are approximately linear as is indicated in fig. 5, where the crosses represent the plot of b against f, the circles A. It will readily be seen that

$$b = 7000 - 3200f,$$

$$A = 7 \cdot 40 - 1 \cdot 60f,$$

or

$$\log p = \log f + 0 \cdot 71 - 1 \cdot 6f - \frac{7000 - 3200f}{T}, \qquad (7)$$

for all films over the range $f = 0$ to $f = 1$, i.e. up to $\sigma = 2{\cdot}7 \times 10^{14}$ atoms per unit of real area. Over this range equation (7) may be used to calculate the spreading force F by means of the Gibbs adsorption isotherm. This equation is best used in its integral form (Bosworth 1932)

$$F = 2{\cdot}303kT\int \sigma d \log p, \tag{8}$$

which on substituting in equation (7) becomes

$$F = 2{\cdot}303k\sigma_m\int fTd\log f + \int 3200 f df(1-0{\cdot}0004T),$$

which on integrating, inserting numerical values and remembering that $F = 0$ when $f = 0$, becomes

$$F = 0{\cdot}037Tf + 140f^2(1-0{\cdot}0004T)\,\text{dynes/cm}. \tag{9}$$

This expression for the spreading force may be split into two parts:

$$F = F' + F'',$$

where F' is the kinetic energy per unit area possessed by the adsorbed particles and F'' is their potential energy per unit area, and

$$F' = \sigma kT, \tag{12}$$

which means that the adions have been treated as an ideal gas in two dimensions. For the corresponding caesium films Langmuir (1932) has made use of the expression

$$F = \sigma_1 kT\frac{\theta}{(1-\theta)}.$$

This equation, however, infers that the condensation coefficient for caesium is proportional to θ, but we have other evidence that every incident atom on the surface adheres no matter what the θ.

The second term in the expression for F is

$$F'' = 140f^2(1-0{\cdot}0004T), \tag{10}$$

and this is to be regarded as the potential energy per unit area possessed by the adions in virtue of their mutual repulsion. The electrostatic energy possessed by an array of dipoles, such as these, has been calculated by Topping (1927) to be

$$F'' = 4{\cdot}51\mu^2\sigma^{5/2} \tag{11}$$

per unit area. Both μ and σ are known and we are therefore in a position to compare the value of F'' calculated by the Gibbs and by the Topping methods. Table II gives the comparison.

TABLE II

f	F'' Gibbs dynes/cm.	F'' Topping dynes/cm.
0·02	0·06	0·15
0·1	1·4	3·6
0·2	5·6	13·0
0·4	22·0	42·0
0·8	90·0	110·0
1·0	140·0	150·0

This table shows that the electrostatic forces are capable of accounting for all the potential energy of the film, and when the Topping method is modified to take into account the interaction between the lattice points on the metal surface and the adions, as well as the mutual interaction of the latter, a better comparison should be obtained. As it is one must conclude that the forces acting between the sodium adions are almost entirely electrostatic in origin. The difficulty that F' calculated from equation (10) has a negative temperature coefficient, while from equation (11) F is to be expected to have a positive temperature coefficient (μ has a positive temperature coefficient), does not prove to be insuperable if the picture of the increase of μ with T adopted above be employed. In such a picture the temperature effect on μ is shown diagrammatically in fig. 6. At low temperatures the dipoles lie in a plane, but at high temperatures a small fraction of adions is lifted out of the plane of normal dipoles. Such a change, although it increases the average μ, does not increase the energy per unit area, rather must it produce a second-order decrease in F, for work must be obtained from the surface by allowing occasional ions to be pushed out of the surface plane. It is permissible then only to compare equations (10) and (11) at low temperatures, and under such conditions the table shows them to be consistent.

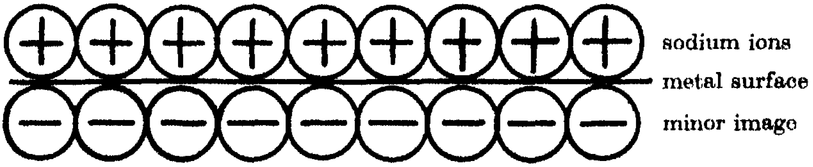
One further point deserves notice. The evaporation energy for dilute films (1·4 e-volts) is much less than the positive-ion work function for sodium (2·5 e-volts) (Bosworth 1935), which is, of course, proof that the sodium is evaporating from the surface mainly as atoms.

We can thus determine the Born cycle for the surface ionization potential $U_i^{(s)}$ of sodium on tungsten in dilute films, i.e. solve the equation

$$U_a + U_i = U_p + \phi + U_i^{(s)},$$

where U_a and U_p are respectively the evaporation energy for atoms and ions from the surface, in the present instance 1·4 and 2·5 e-volts respectively, U_i is the ionization potential of sodium = 5·1 V, and ϕ is the work function of tungsten (4·5 V). $U_i^{(s)}$ is accordingly $-0{\cdot}5$ V; thus a dilute film of sodium on tungsten is adsorbed as ions, bringing sodium into line with the other alkalis, where since a beam of atoms incident on a hot filament is almost quantitatively reflected as ions is to be taken as an indication for ionic adsorption. Thus when a sodium atom is brought up to a clean tungsten surface, ionization occurs and the resultant ion is adsorbed by electrostatic forces, the reaction

A. Surface covered with normal adsorbed ions.



B. Same surface with some of the adions activated to a higher level.

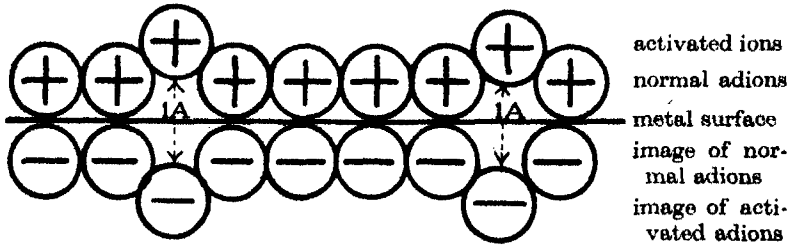


FIG. 6

$$\mathrm{Na} + \mathrm{W} \rightarrow \mathrm{Na}^{+} + \mathrm{W}^{-}$$

(wherein W is to be understood as the tungsten surface) is exothermic in amount 0·5 e-volt.

In the film of minimum work function however,

$$\phi = 1{\cdot}7, \quad U_a = 0{\cdot}7, \quad \text{and} \quad U_p = 1{\cdot}9,$$

so that $U_i^{(s)} = +2{\cdot}2$ V, or the mutual interaction of the sodium ions has raised the ionization potential on the surface by 2·7 V.

EVIDENCE FOR A TWO-PHASE SYSTEM ON THE SURFACE AT CERTAIN CONCENTRATIONS

On substituting the value for the surface potential V (in c.g.s. units) in the Topping equation, the following expression for F'' is obtained:

$$F'' = 0{\cdot}114\,V^2\sigma^{\frac{1}{2}}.$$

Reference to fig. 5 in the previous paper shows that above $f = 1$ not only does V decrease with rising σ but also over a certain range $V^2\,\sigma^{\frac{1}{2}}$ decreases or F'' decreases until at about $f = 2$ V becomes sensibly constant and F'' again rises with σ. This means that over this range of concentrations the films tend to be unstable, for if the film could break up into a concentrated phase ($f \simeq 2$) and a dilute film, a nett decrease in F'' (the surface free energy) would be effected. A similar conclusion is drawn from a consideration of the heat of evaporation. For very concentrated films the heat of evaporation must be the same as from bulk sodium, viz. 25,000 cal.; but for all the films more concentrated than $f = 0{\cdot}5$ which have been studied above, this heat is less than 25,000 cal., and consequently there must exist a temperature below which any such film cannot be stable but must split up into a dilute phase for which $f \leqslant 0{\cdot}5$ and a concentrated phase for which f is about 2. Exactly how concentrated, however, will not be determined until the heats of evaporation are determined for values of f from 1 to ∞. Evidence for the existence of such two-phase systems has also been obtained from an examination of the Schottky effect in these films (Bosworth 1937). It is interesting to note at this stage that the electron work function of the dilute film with the same heat of evaporation as bulk sodium is sensibly the same as the photoelectrically determined work function for bulk sodium, viz. 2·5 e-volts. Much the same figure appears also to be obtained from contact potential measurements on the final film formed when sodium is deposited on tungsten at a low temperature. From deposits formed at higher temperatures ($> 350°$ K.) such a low final value is not attained, presumably because in such films the surface at a definite θ comes to equilibrium with the vapour.

Let us consider, for example, the equilibrium between the film of minimum work function

$$\log p = -0{\cdot}89 - \frac{3750}{T}$$

and bulk sodium

$$\log p = 7{\cdot}68 - \frac{5410}{T}.$$

The vapour pressure of the two systems will be equal when

$$T = 182^\circ \text{K.},$$

and at all lower temperatures this film must be unstable.

In the same way at 122° K. a film of concentration $f = 0{\cdot}8$ will be in equilibrium with solid sodium. These figures are, however, only qualitatively correct, for a very little change in the evaporation energy will profoundly effect these temperatures, and the points used for the calculation of the heats of evaporation in fig. 4 have already been obtained from smoothed curves in fig. 3. The computations above indicate only the existence and the order of these critical temperatures and the fact that they increase as f increases.

One further point deserves notice. A concentrated film ($f = 1{\cdot}5$) behaves in a quite different way on heating to more dilute films. One would expect as evaporation proceeded that the surface potential would first rise and then fall as one moved in the reverse direction along the V-σ curve. No such change has, however, ever been observed. When evaporation occurs it proceeds always as a drop in the surface potential. This drop may be made quite slow by making the evaporating temperature low enough, and no positive change has ever been observed. If the temperature is too low for evaporation the only thermal effect appears to be a small negative change completely reversible on cooling again. We can only assume that for films of this order of thickness aggregates of condensed phase tend to be built up in equilibrium with patches of an expanded phase ($f \simeq 0{\cdot}5$), and when evaporation proceeds it goes by first a two-dimensional isopiestric evaporation of the condensed phase followed by the normal evaporation of the disperse phase. This condensed phase appears to be complete just beyond $f = 2{\cdot}0$, i.e. $\sigma = 5{\cdot}4 \times 10^{14}$, for at about this point the Topping spreading force begins to rise again on further increasing σ, and it is not without interest to note that if we accept the figure given by Simon and Vohsen (1928) for the lattice constant of sodium (viz. 4·24 A), and pack spheres of this diameter on to a surface in a hexagonal lattice regardless of the underlying structure, at a surface concentration of $8{\cdot}6 \times 10^{14}$ atoms per sq. cm., the spheres will just be touching. If now we assume that the total spreading force is a sum of the Topping force and a kinetic term of the type

$$F'(A - A_0) = kT, \tag{13}^*$$

---

[*] The difference between equations (13) and (12) is well within experimental error over the range of σ studied in the evaporation experiments.

where $A_0 = \frac{10^{-14}}{8 \cdot 6} = 11 \cdot 7$ A is so chosen that the adions are prevented from packing together closer than $8 \cdot 6 \times 10^{14}$ per sq. cm., i.e. if we write

$$F = 0 \cdot 114 V^2 \sigma^{\frac{1}{2}} + \frac{kT}{\left(\frac{1}{\sigma} - \frac{1}{\sigma_0}\right)}, \tag{14}$$

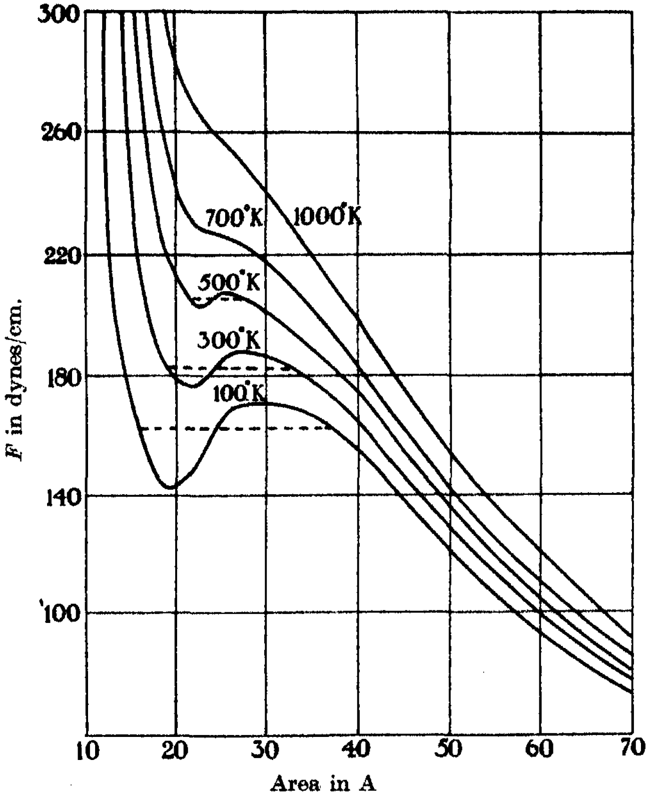


FIG. 7

we are in a position to draw *F*-*A* curves for the system. Fig. 7 shows five such curves at the fixed temperatures of 100, 300, 500, 700 and 1000° K. The curves show all the characteristics of a two-phase system with a critical temperature of about 700° K. No absolute accuracy can be claimed for this figure, for it would be profoundly modified by choosing different A's, and further, no attempt has been made to allow for the temperature variation of the Topping force. The figure, however, illustrates the type of system we are dealing with, and shows that below a certain temperature the system will at a certain concentration ($\sigma = 2{\cdot}7 \times 10^{14}$ at 100° K. on the curves given) split into a concentrated and dilute phase. We are therefore led to expect that the true spreading forces, if they could be measured (and not calculated from averaged properties of the surface), would just above $f = 1$ become independent of the surface concentration over a range of σ's. This point, too, it is hoped to test with the three-filament apparatus.

SUMMARY

Experimental conditions to be observed when studying temperature effects in films by the method of contact potentials are discussed, and the behaviour of films of sodium on heating examined by these means.

It is found that on heating, an adion in the film may be first excited to a higher energy level while still remaining essentially an adsorbed ion. In the higher energy state the ion is farther from the surface than is the normal state, and consequently its effective dipole moment and contribution to the surface potential is increased.

On further heating the sodium ion evaporates (as an atom) and the evaporation rate falls off rapidly as the film becomes more dilute. Curves are given showing how the rate of evaporation varies with θ and with temperature, and from those curves the spreading forces, in the Gibbs sense, have been calculated.

This spreading force proves to have a small negative temperature coefficient, and its value at low and ordinary temperatures compares so well with the electrostatic energy per unit area of the surface that one may conclude that the forces between the ions on the surface are purely electrostatic.

A calculation of the surface ionization potential for dilute films shows that this must be negative or that a sodium atom brought up to a bare tungsten surface will, like the more electropositive alkalis, be adsorbed as an ion.

From the behaviour of the spreading forces and rates of evaporation as the surface concentration increases it is deduced that concentrated films will have a tendency to split into a condensed and a diffuse phase, and this tendency explains the peculiarities in the evaporation of such films.

A short treatment of the thermal effects on the surface potential of tungsten itself is also given.

REFERENCES

Becker, J. A. 1932 *Trans. Faraday Soc.* **28**, 148–58.
Becker, J. A. and Brattain, W. H. 1934 *Phys. Rev.* **45**, 694–705.
de Boer, J. H. and Veenemans, C. F. 1935 *Physica*, **2**, 529–34.
Bosworth, R. C. L. 1932 *Trans. Faraday Soc.* **28**, 903–12.
— 1935 *Proc. Roy. Soc.* A, **150**, 58–76.
— 1937 *Trans. Faraday Soc.* **33**, 590–6.
Bosworth, R. C. L. and Rideal, E. K. 1937 *Proc. Roy. Soc.* A, **162**, 1–31.
Killian, T. J. 1926 *Phys. Rev.* **27**, 578–87.
Kuhn, F. 1932 *Ann. Phys. Lpz.*, **15**, 825–41.
Langmuir, I. 1932 *J. Amer. Chem. Soc.* **54**, 2798–832.
Langmuir, I. and Kingdon, K. H. 1925 *Proc. Roy. Soc.* A, **107**, 61–79.
Lennard-Jones, J. E. 1932 *Trans. Faraday Soc.* **28**, 333–59.
Richardson, O. W. 1916 "The Emission of Electricity for Hot Bodies", pp. 39–42.
Simon, F. and Vohsen, E. 1928 *Z. phys. Chem.* **133**, 165–87.
Topping, J. 1927 *Proc. Roy. Soc.* A, **114**, 67–72.

The Electronic Specific Heat and X-Ray Absorption of Metals, and Some Other Properties Related to Electron Bands

BY H. JONES AND N. F. MOTT, F.R.S., *H. H. Wills Physical Laboratory, University of Bristol*

(*Received* 3 *June* 1937)

According to the electron theory of metals the energies of the possible stationary states of the electrons in a metal form a continuum; it is usual to denote by $N(E)\,dE$ the number of such states with energies between E and $E+dE$ per unit volume. In a metal at the absolute zero of temperature all states are occupied up to a maximum energy $E = \zeta_0$.

It is the purpose of this note to discuss the quantity $N(E)$ in connexion with the crystal structure, specific heat and X-ray absorption spectra of metals.

If the crystalline field had no influence on the metallic electrons, we should have (Mott and Jones 1936)

$$N(E) = \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} E^{\frac{1}{2}}. \tag{1}$$

We refer to this as the "free electron" formula.

In a real metal, however, the possible stationary states are divided into zones; the density of states for the first zone and part of the second zone are shown in fig. 1 for the body-centred cubic structure. (In this figure the part of the curve *OABC* is from actual calculations by Jones (1937) for α and β brass, the rest is a rough estimate, making use of the theorem that $N(E) \propto \sqrt{E}$ at the bottom of a zone, and $N(E) \propto \sqrt{(E_0 - E)}$ at the highest point E_0 of a zone.)

A kink of the type *ABC* occurs in the $N(E)$ curve for energies corresponding to every set of reflecting planes of the lattice as was first shown by Kronig (1931). The energy of the point *B* will be, for cubic structures, about

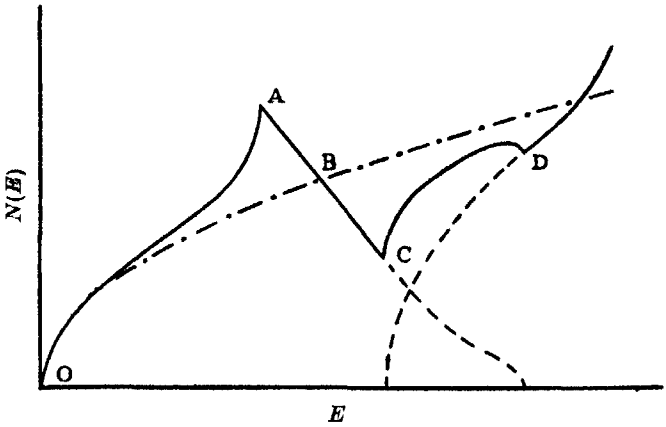


FIG. 1— — · — · — for free electrons; - - - - - individual zones; ———— total $N(E)$.

$$E = \frac{h^2}{8ma^2}(n_1^2 + n_2^2 + n_3^2), \tag{2}$$

where a is the side of the unit cell and $n_1 n_2 n_3$ are the Miller indices of the set of planes considered. If, as is usual in the electron theory of metals, we represent each electronic state by a point in k-space, and if the lines in fig. 2 represent the planes corresponding to the indices $(n_1 n_2 n_3)$ across which there is an energy discontinuity, then *A* in fig. 1 corresponds to the curved lines in fig. 2, representing energy surfaces which *just* touch the planes of energy discontinuity. The kink at *A* is due to states near *P*, the sudden rise at *C* to states near *Q*.

In Table I we give, for the body-centred and face-centred cubic structures, corresponding to each set of planes:

(*I*) The structure factor.

(*II*) The energy of a free electron with energy just great enough to be reflected by these planes. This is given by formula (2). We give the quantity

$$\text{Energy} \times (\text{Atomic volume})^{\frac{2}{3}}$$

in e-volts, the atomic volume being the volume per gram-atom.

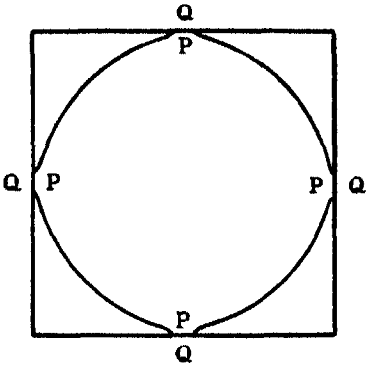


Fig. 2

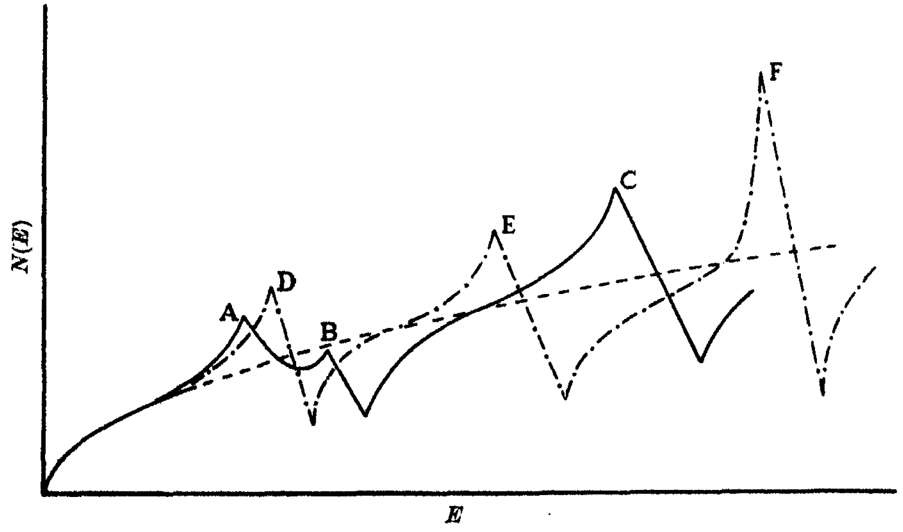


Fig. 3—Density of states in body- and face-centred cubic metals. - - - - - - - - - free electrons; ———— face-centred cubic lattice; — · — · — body-centred cubic lattice. The peaks *A*, *B*, *C* arise from the planes (111), (200), (220) respectively, and *D*, *E*, *F* from the planes (110), (200), (211).

(*III*) The number of electrons which will fill all states up to that of energy given by (2), again for free electrons. This is given by

$$N = \frac{8\pi}{3}\left(\frac{2mE}{h^2}\right)^{\frac{3}{2}}. \tag{3}$$

In order to construct the $N(E)$ curve we must draw a kink of the type ABC (fig. 1) for each plane where the structure factor does not vanish. We assume that the energy given in Table I coincides with the middle of the kink (point B, fig. 1). Fig. 3 shows the type of curve obtained.

TABLE I

		110	111	200	210	211	220
Indices of plane		110	111	200	210	211	220
Structure factor	*b–c*	2	0	2	0	2	2
	f–c	0	4	4	0	0	4
Energy × (atomic volume)$^{\frac{2}{3}}$	*b–c*	47·2	—	94·4	—	141·6	—
	f–c	—	44·6	59·5	—	—	119·0
Number of electrons	*b–c*	1·48	—	4·21	—	7·7	—
	f–c	—	1·36	2·10	—	—	5·9

SPECIFIC HEAT

The heat capacity per gram-atom of the *electrons* in any metal is given by

$$C_v = \tfrac{2}{3}\pi^2 k^2 N(\zeta_0)\, T \times \text{atomic volume}. \tag{4}$$

It is convenient to compare (4) with the Sommerfeld formula, obtained by substituting (1) for $N(E)$; this gives

$$(C_v)_{\text{Somm.}} = 3{\cdot}26 \times 10^{-5}\, n_0^{\frac{1}{3}} \Omega_0^{\frac{2}{3}} T, \tag{4·1}$$

where Ω_0 is the volume per gram-atom, and n_0 the number of electrons per atom outside a closed shell.

We see from fig. 3 that $N(E)$, for a given structure, may be greater or less than the Sommerfeld value, so that, at first sight, for any random selection of metals with various n_0, we should expect C_v to be sometimes greater and sometimes smaller than the Sommerfeld value (4·1). On the other hand, the investigations of Jones on the crystal structure of metals and alloys show that the crystal structure with lowest energy is in general that with lowest ζ_0. *This will in general be that in which* $N(\zeta_0)$ *is as large as possible.*

Fig. 3 suggests that for one or other of the two cubic structures, $N(E)$ will be greater than the Sommerfeld value (4·1) for nearly all values of n_0. Since metals have several other structures to choose from, it is probable that for metals having their most stable structure, $N(E)$ is *always* greater than the Sommerfeld value.

In Table II we give the electronic specific heats for all metals for which they have been measured, extrapolated to 1° K.:

TABLE II—ATOMIC HEAT AT 1° K. $\times 10^{-4}$ CAL./DEGREE/GRAM-ATOM

Metal	n_0	Calc. (4·1)	Observed	Ref.
Cu	1	1·20	1·78	(1)
Ag	1	1·54	1·6	(1)
Zn	2	1·80	2·33–2·97	(3)
Hg	2	2·35	3·75	(3)
Tl	3	3·06	3·8	(3)
Sn	4	3·15	3·46	(3)
Pb	4	3·54	7·07	(3)
Ta	—	—	27	(3)
Nb	—	—	(70)	(3)
Ni	—	—	17·4	(1)
Pd	—	—	31	(2)
Pt	—	—	16·1	(1)

(1) Observed, Keesom and collaborators (1934), (1935), (1937).
(2) Observed, Pickard (1937).
(3) Daunt and Mendelssohn (1937), from magnetic threshold curves; the values apply of course to the non-supraconducting state.

We see, as we expect, that in all cases the electronic specific heat is greater than the Sommerfeld value.

For copper, detailed calculations (Jones 1937) of $N(E)$ have already been made with a view to determining the phase diagram of Cu-Zn alloys. From these calculations we obtain $N(\zeta_0) = 0{\cdot}310$ electron per atom per e-volt compared with 0·219 given by Sommerfeld's formula (1). The corresponding value,

$$(C_v)_{\text{Theoretical}} = 1{\cdot}66 \times 10^{-4}\, T \text{ cal./degree/gram-atom},$$

compares well with the observed value.

The high specific heat of Ni, Pd and Pt has already been discussed (Mott 1935) and attributed to the *vacant places* in the *d*-shells of these transition metals.

Tantalum is also a transition metal and we must attribute its high specific heat to the same cause, namely, an incomplete *d*-shell. With a view to correlating the properties of this and other transition metals we have made a calculation of the density of states in the *d*-band, which we discuss in the next section.

DENSITY OF STATES IN d-BAND

The number of states in the d-band is 10 per atom. As first shown by Bethe (1929) the d-level is split in a cubic field into two levels having weights 6 and 4. Thus in the metal we shall have two d-bands, which may or may not overlap. The number of states per atom in the two bands will be 6 and 4. The d-band with weight 6 has wave functions having the form $xyf(r)$; that with 4, $(x^2-y^2)f(r)$. It is important to know which band is the lower. Let us first take the body-centred cubic structure. Then any one atom is surrounded by eight neighbours, so that the potential energy of an electron is lowest in the (111) directions. Now the wave functions $xyf(r)$ are greatest in these directions, while the function $(x^2-y^2)f(r)$ is zero. Hence it follows that the xy band is the lower.

Hence for the body-centred cubic structure the band with 6 electrons lies below that with 4.

For the face-centred structure the position is reversed; the band with 6 is the higher.

We have been able, using the approximation of tight binding, to calculate $N_d(E)$ for the lower d-band for the body-centred cubic structure. The three atomic d-wave functions which combine to form the lower in energy of the two parts of the d-band in body-centred cubic crystals are of the form

$$\phi_1 = xyf(r), \quad \phi_2 = yzf(r), \quad \phi_3 = zxf(r).$$

We consider a crystal of N atoms (N being a very large number) and form the following three functions:

$$\psi_{n\mathbf{k}} = \frac{1}{\sqrt{N}}\sum_{l} e^{ia(\mathbf{l},\mathbf{k})}\,\phi_n(\mathbf{r}-a\mathbf{l}), \quad n = 1, 2, 3, \tag{5}$$

where the summation is over the N values of $\mathbf{l}$ which correspond to the position of each atom in the lattice. a is the lattice constant of the body-centred cubic crystal. With the three functions (5) we can form the secular equation for the energies in the lower part of the d-band. This is as follows:

$$\begin{vmatrix} H_1-E & H_{12} & H_{13} \\ H_{21} & H_1-E & H_{23} \\ H_{31} & H_{32} & H_1-E \end{vmatrix} = 0, \tag{6}$$

where
$$
\begin{aligned}
H_1 &= 8A\cos\xi\cos\eta\cos\zeta,\\
H_{12} = H_{21} &= -8B\sin\xi\cos\eta\sin\zeta,\\
H_{23} = H_{32} &= -8B\sin\xi\sin\eta\cos\zeta,\\
H_{13} = H_{31} &= -8B\cos\xi\sin\eta\sin\zeta,
\end{aligned}
$$
$$\xi = \tfrac{1}{2}ak_x, \quad \eta = \tfrac{1}{2}ak_y, \quad \zeta = \tfrac{1}{2}ak_z.$$

The parameters A and B are given by

$$A = \int \phi_1(xyz),\ H\phi_1(x+\tfrac{1}{2}a,\ y+\tfrac{1}{2}a,\ z+\tfrac{1}{2}a)\,d\tau,$$

$$B = \int \phi_1(xyz),\ H\phi_2(x+\tfrac{1}{2}a,\ y+\tfrac{1}{2}a,\ z+\tfrac{1}{2}a)\,d\tau.$$

The operator H is essentially the Hamiltonian operator (cf. Mott and Jones 1936, p. 66). For very tight binding, i.e. when the spread of the atomic wave functions is small compared with the interatomic distance, the parameters A and B become approximately equal. To make the calculation manageable and yet to obtain the essential features of the $N(E)$ curve for the d-band, we make the assumption $A = B$.

There are several important simplifications to be noted. First it is only necessary to calculate the energy surfaces in a single octant. This can be seen as follows:

Let
$$\epsilon = (E/8B) - (A/8B)\cos\xi\cos\eta\cos\zeta$$
and
$$
\begin{aligned}
a &= \sin\xi\cos\eta\sin\zeta,\\
b &= \cos\xi\sin\eta\sin\zeta,\\
c &= \sin\xi\sin\eta\cos\zeta;
\end{aligned}
$$
the secular equation (6) then becomes
$$\epsilon^3 - (a^2+b^2+c^2)\,\epsilon + 2abc = 0, \tag{7}$$
and it can now be seen that this equation is invariant with respect to reflexions in any of the coordinate planes, although a, b and c themselves are not.

The second simplification is due to the fact that abc is always small inside the zone; in fact the product vanishes over each of the planes $\xi = 0$, $\eta = 0$, $\zeta = 0$, $\xi = \tfrac{1}{2}\pi$, $\eta = \tfrac{1}{2}\pi$, $\zeta = \tfrac{1}{2}\pi$. It follows therefore that one solution of (7) is approximately $\epsilon = 0$. The energy surfaces corresponding to $\epsilon = 0$ have exactly the same form as the energy surfaces for the s-band. Estimations of the differences between the actual solution and $\epsilon = 0$ are found to be unimportant. The $N(E)$ curve corresponding to one solution of (7) will therefore be like that already given for the s-band (cf. Mott and

Jones 1935, p. 85, fig. 38 (*b*)), the characteristic feature of which was the sharp peak at the centre. The third simplification is that the remaining two solutions of (7) are essentially the same.

The $N(E)$ curves corresponding to these two solutions are mirror images of each other in the mid-point of the band (see fig. 5).

Finally, with the assumption $A = B$ the energy surfaces for the one solution which it is necessary to find are relatively simple. For instance, consider the plane $\zeta = 0$; then $a = b = 0$; and we find from (7)

$$\epsilon = \pm \sin\xi \sin\eta$$

and therefore $$E/8A = \cos(\xi \pm \eta).$$

The $N(E)$ curve for the solution considered has a sharp peak at the end of the band. Fig. 4 shows the form of the zone in k-space. The energies at various points within this zone are easily obtained from (7). We find, for example, that $E/8A$ is equal to -1 at the origin, 0 at the point C, and has its maximum value $+1$ along the lines AB, AD, etc. The energy surfaces for values of $E/8A$ slightly less than $+1$ are groove-like surfaces about these lines. The energy surfaces near the origin are closed surfaces. The $N(E)$ curve was obtained by approximating to the actual surfaces by simple ones, spheres, cylinders, etc., and obtaining the number of states enclosed within specified ranges of E. Fig. 5 is the sum of the three $N(E)$ curves corresponding to the three solutions of (7). In fig. 5 the $N(E)$ curve drops suddenly from a large value to zero at the boundaries of the band; this is due to the assumption $A = B$ which was made merely in the interests of simplicity. In fact, A is always greater than B and this results in the gradual fall of $N(E)$ to zero at the ends of the band as indicated by the dotted curves.*

PROPERTIES WHICH DEPEND ON THE *d*-BAND

In metals for which $N_d(\zeta_0) \gg N_s(\zeta_0)$, $N_d(\zeta_0)$ is important for the following properties:

(1) Specific heat due to electrons (see above).

(2) Magnetic susceptibilities; large $N_d(\zeta_0)$ will in general mean large susceptibility, though the two will not in general be proportional, owing to "exchange forces" (Mott and Jones 1936, p. 189).

---

* Slater (1936) has given curves showing the density of states in the *d*-band for a face-centred cubic metal (copper). Slater uses the method of Wigner and Seitz. We believe that for the *d*-band, in which the atoms overlap but little, the method of "tightly bound electrons" used here should be more accurate.

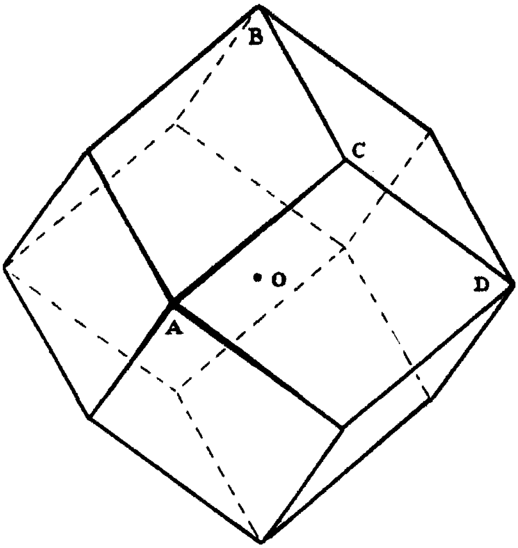


Fig. 4

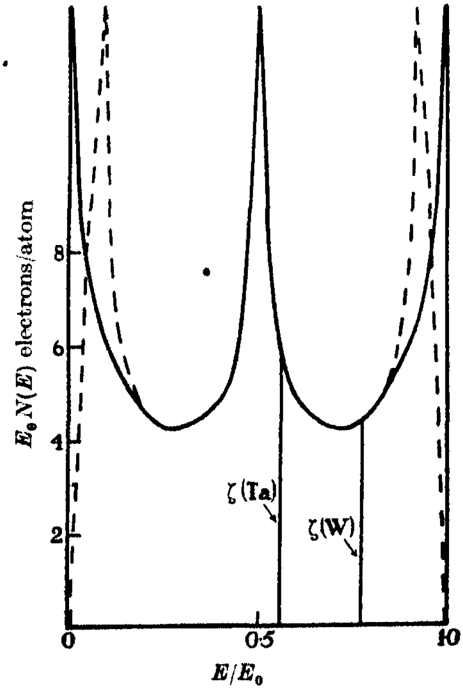


Fig. 5—The density of states for the lower d-band in body-centred cubic metals. Full line, using approximations of this paper; dotted line, correct form.

(3) Resistance at high temperatures; it has been shown by one of us that if $f_0(E)$ is the Fermi distribution function

$$\frac{R}{T} = \text{const.} \Big/ \int_0^\infty \frac{1}{N_d(E)} \frac{\partial f_0}{\partial E} dE. \tag{8}$$

Since $\partial f_0/\partial E$ vanishes except in a range of order kT about $E = \zeta_0$, it follows that if in this range $N_d(E)$ is concave to the E-axis, then R/T is concave to the T-axis, and vice versa (Mott and Jones 1936, p. 268).

(4) Integrated thermoelectric power. This is roughly proportional to*

$$(dN_d/dE)_{E=\zeta_0}.$$

In the following table we give the relevant data for the transition metals which have the body-centred cubic structure.

TABLE III—TRANSITION METALS HAVING BODY-CENTRED CUBIC STRUCTURE

Metal	Fe	(Mn)	Cr	V	Mo	Nb	W	Ta
Electrons outside closed p-shell	8	(7)	6	5	6	5	6	5
Susceptibility × 10^6	—	(489)	151	76	54	121	40	145
Thermoelectric power S at 100° C.	—	—	—	—	+8·0	—	+3·6	−6·7
Electronic specific heat × 10^{-4}	—	—	—	—	—	(70)	—	27
Resistance at high temperatures							Convex to T-axis	Concave to T-axis

Now in order to estimate $N_d(\zeta)$ we must know how many electrons are in the d-band, and thus how many are in the s-band. In alloys whose crystal structure is determined by the Hume-Rothery rule the body-centred β phase occurs when there are about 1·5 valency electrons per atom. We shall tentatively assume that there are 1·5 s-electrons per atom in these metals. This leaves 3·5 and 4·5 per atom in the d-band for Ta and W respectively. With these numbers the d-band is full up to the points shown in fig. 5. It will be seen that the differences between Ta and W can then be accounted for qualitatively, as follows: since $N_d(E)$ is greater for Ta than for W, the susceptibility is greater for Ta. Since $d(\log N)/dE$ is large and negative for Ta, small and positive for W, so are the thermoelectric powers S. The high electronic specific heat of Ta is accounted for by the peak in the $N(E)$ curve. Finally, since $N(E)$ is convex to the E axis, so is the R/T curve for W. For Ta on the other hand, the $N(E)$ curve, though actually convex to the axis at the point $E = \zeta_0$, is sharply concavė over a small range round $E = \zeta_0$ including the peak. We have calculated the values of R/R_0 from formula (8). A half-width of the band of 0·47 e-volt was found to give the best agreement with experiment which could be obtained with the given $N(E)$ curve. The following values were obtained:

---

* Mott and Jones (1936), p. 310, equation (116). Since $S \propto -\partial(\log \sigma)/\partial E$ and $\sigma \propto 1/N_d(E)$, our result follows.

T (°K.)	273	546	1092
R/R_0 calc.	1·0	1·89	3·58
R/R_0 obs.	1·0	2·00	3·54

The comparison between tungsten and tantalum given here depends on the assumption that the number of electrons per atom in the s-band is 1·5; our results are also valid if the number is greater than 1·5, but not if much less. Ta is superconducting; it is interesting that all superconducting metals with full d-shells have more than one conduction electron per atom. Ni, Pd and Pt, which are not supraconductors, have only 0·6.

X-RAY EMISSION AND ABSORPTION

Kronig (1931) was the first to account for the fine structure shown by the X-ray absorption edges of solids in terms of the $N(E)$ function. Recent experimental work in the ultra soft region (Skinner and Johnston 1937) justifies a closer analysis of the fine structure to be expected.

The intensity of X-ray absorption may be written

$$I(E) = N(E)p(E),$$

where $N(E)$ is the density of states as before and $p(E)$ the transition probability. The same formula has been used in discussing the X-ray emission (Jones, Mott and Skinner 1934). $p(E)$ will be proportional to the square of a matrix element of the type

$$\int \psi_f x \psi_i \, d\tau, \tag{9}$$

where ψ_i is the wave function of the initial X-ray level and ψ_f of the final level in the conduction band. We may expand ψ_f in a series of spherical harmonics

$$\psi_f = \Sigma A_n f_n(r) \, S_n(\theta, \phi),$$

where the functions S_n are spherical harmonics of degree n. The functions $f_n(r)$ may be normalized in any convenient way so that

$$\Sigma \mid A_n \mid^2 = 1.$$

Owing to the small extension of the function $\psi_i(r)$ corresponding to the X-ray level, the transition probabilities will depend on the form of ψ_f near the nucleus, where f_0, f_1 etc., will be very similar to the radial wave functions of the free atom in the s, p, ..., states. The variation in the transition probability near an absorption edge will thus depend mainly on the coefficients A_n. In particular, if the initial level is an s-level (e.g. K, L_{I}), then $p(E)$ is proportional to $\mid A_1 \mid^2$; if it is a p-level (e.g. L_{II}, L_{III}) it depends on A_0 and A_2, but owing to the small values of the matrix elements (9) for $p \rightarrow d$ transitions, we may neglect A_2 except for transition metals, where the density of d-states is abnormally large.

It is convenient to define a "density of s-states" N_s and a "density of p-states" N_p given by

$$N_s(E) = N(E) \mid A_0 \mid^2,$$

$$N_p(E) = N(E) \mid A_1 \mid^2.$$

Then if the X-ray level is an s-level,

$$I(E) \propto N_p(E),$$

and if the X-ray level is a p-level,

$$I(E) \propto N_s(E).$$

In order to interpret the X-ray absorption spectra we need to know the characteristics of N_s, N_p.

Many of the details may be deduced at once from the investigations of Jones, Mott and Skinner (1934); thus:

(*a*) At the bottom of the first zone

$$N_s \propto E^{\frac{1}{2}}, \quad N_p \propto E^{\frac{3}{2}}. \tag{10}$$

(*b*) At the bottom of any other zone there is no generally valid rule. In the limiting case of high energies, in which Bloch's approximation of nearly free electrons is satisfactory at the bottom of the second zone,

$$N_s \propto E^{\frac{3}{2}}, \quad N_p \propto E^{\frac{1}{2}}.$$

This arrangement for the first and second zones we call the normal case. For certain metals, Jones, Mott and Skinner (1934) showed that an inversion may occur so that at the beginning of the second zone

$$N_s \propto E^{3/2}, \quad N_p \propto E^{1/2} \text{ (inverted case).}$$

This actually happens for Mg and Li.

(*c*) The sharp peak (fig. 1) will not occur for both N_s and N_p. In the "normal" case N_s shows the peak, in the inverted case N_p.

The types of curve obtained near the energies corresponding to any set of Miller indices (formula (2)) are thus as in fig. 6.

(*d*) At the head of a band no general prediction is possible, though $N(E)$ is proportional to $\sqrt{(E_0 - E)}$, where E_0 is the energy at the head of the band.

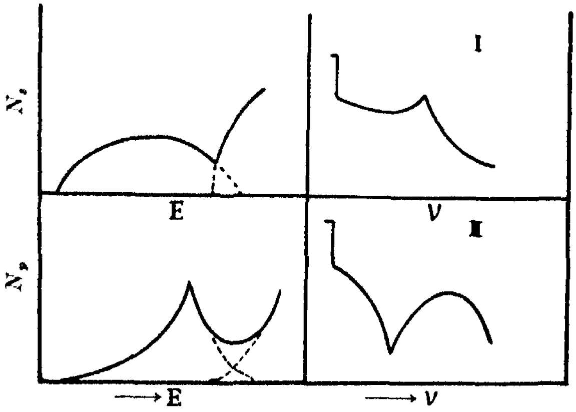


FIG. 6—*Left,* density of *s*- and *p*-states, dotted lines for separate zones, full lines total. *Right,* intensity of transmitted radiation near absorption edge. Type I will be obtained for absorption from *p*-levels (L_{II}, L_{III}) sufficiently far from the edge; Type II from *s*-levels (K, L_I). Near the edge the behaviour *may* be reversed.

SUMMARY

The energy bands of metals are discussed in relation to the following physical properties:

(1) The electron specific heat of metals having closed *d*-shells. It is shown that the linear term is in general greater than that given by Sommerfeld's free electron formula.

(2) Electronic specific heat, magnetic and electrical properties of transition metals. These are related to the shape of the *d*-band, which is worked out for body-centred cubic metals, e.g. Ta and W and α-Fe. It is shown that Ta, like other supraconductors, has more than one "conduction" electron per atom.

(3) The fine structure of the X-ray absorption edge.

REFERENCES

Bethe 1929 *Ann. Phys., Lpz.*, **3**, 133.
Daunt and Mendelssohn 1937 *Proc. Roy. Soc.* A, **160**, 127.
Jones 1937 *Proc. Phys. Soc.* **49**, 50.
Jones, Mott and Skinner 1934 *Phys. Rev.* **45**, 378.
Keesom and Clark 1935 *Physica, 'sGrav.*, **2**, 513.
Keesom and Kok 1934 *Physica, 'sGrav.*, **1**, 175.
— — 1936 *Physica, 'sGrav.*, **3**, 1035.
Kronig 1931 *Z Phys.* **70**, 317.
Mott 1935 *Proc. Phys. Soc.* **47**, 571.
Mott and Jones 1936 "The Theory of the Properties of Metals and Alloys." Oxford.
Pickard 1936 *Nature, Lond.*, **138**, 123.
Slater 1936 *Phys. Rev.* **49**, 537.
Skinner and Johnston 1937 *Proc. Roy. Soc.* A (in the Press).

The Elastic Stability of a Long and Slightly Bent Rectangular Plate under Uniform Shear

By D. M. A. LEGGETT, B.A., *Trinity College, Cambridge*

(*Communicated by R. V. Southwell, F.R.S.—Received* 5 *February* 1937)

1—The problem of the elastic stability of a plane rectangular plate when subjected to uniform shear has been approximately solved for various conditions (Cox 1933; Timoshenko 1936). In the case of an indefinitely long strip an exact solution has been found (Southwell and Skan 1924), but it appears that no attempt has been made to investigate what happens if the plate is no longer plane. It is accordingly the object of this paper to consider the stability of a long strip, slightly curved, when its two side edges are subjected to uniform shear.

2—In what follows we assume that the thickness and curvature of the plate are constant, and that the edges of the plate are two generators and two lines of curvature. It is, moreover, further assumed that the plate is thin as in all similar stability problems, and that it is of such length that the boundary conditions over the two curved ends can be ignored.

We first consider the equilibrium stresses and displacements caused by applying a uniform shearing force to the two straight edges. Referring the system to ordinary cylindrical co-ordinates, it can be shown without difficulty from the equations of equilibrium and the equations connecting up the coefficients of stress and strain (Love 1927, §§ 22, 59, 69) that the set of displacements

$$u_r = 0, \quad u_\theta = 0, \quad u_z = \tau\theta,$$

where τ is constant, can be produced by a stress system in which all the coefficients of stress vanish except $\widehat{\theta z}$ which has the value $\mu\tau/r$.

3—We now confine our attention to the middle surface of the plate and choose a new system of co-ordinates. We take the generator and line of curvature through the centre of the plate as axes ξ, η, so that the edges of the plate are $\xi = \pm a/2$, $\eta = \pm b/2$, and introduce a third axis ζ, normal to the middle surface in its unstrained condition, and such that ξ, η, ζ form a right-handed system. The displacement of any point of the middle surface is u, v, w, with reference to these axes. The radius of curvature of the plate is ρ, its thickness is $2h$, and the external force system consists of a shear resultant S acting uniformly over the two straight edges.

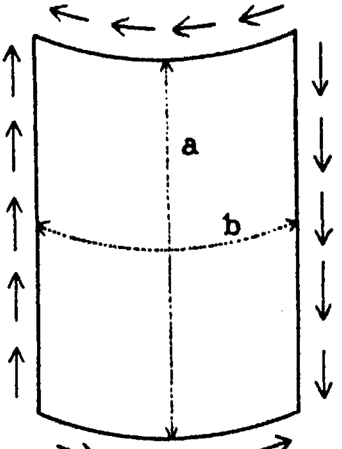


FIG. 1

Since it is supposed that the edges of the plate are free to move in the ordinary equilibrium displacement, it follows from § 2 that the actual displacements are

$$u = -\frac{S\eta(1+\sigma)}{Eh}, \quad v = 0, \quad w = 0.$$

We shall now use the shell equations obtained by Dean for problems of this nature (1925), when second-order terms which are multiplied by $1/\rho$ have been omitted, a legitimate approximation since we are assuming that $1/\rho$ is small. The above set of displacements now satisfies the shell equations exactly, and if the configuration is one of neutral equilibrium, the shell equations can also be satisfied by the displacements

$$u = -\frac{S\eta(1+\sigma)}{Eh} + u', \quad v = v', \quad w = w',$$

where u', v', w' are arbitrarily small but not all zero. Subtracting the shell equations for the two configurations, and ignoring terms of order above the first in u', v', w', we deduce the three fundamental stability equations.

These are long but can be simplified by making two assumptions. First, that $1/\rho$ is small and of the same order as h; and second, that w' is of larger order than u' or v'. The latter assumption follows at once from a consideration of the limiting case when the plate is flat, and from the fact that we are considering a plate which is only slightly bent. As a result of making these approximations and only retaining the terms of the highest order of magnitude, the equations are

$$\frac{\partial}{\partial\xi}\left[\frac{\partial u'}{\partial\xi}+\sigma\left(\frac{\partial v'}{\partial\eta}-\frac{w'}{\rho}\right)\right]+\frac{(1-\sigma)}{2}\frac{\partial}{\partial\eta}\left(\frac{\partial v'}{\partial\xi}+\frac{\partial u'}{\partial\eta}\right)=0, \tag{1}$$

$$\frac{(1-\sigma)}{2}\frac{\partial}{\partial\xi}\left(\frac{\partial v'}{\partial\xi}+\frac{\partial u'}{\partial\eta}\right)+\frac{\partial}{\partial\eta}\left[\left(\frac{\partial v'}{\partial\eta}-\frac{w'}{\rho}\right)+\sigma\frac{\partial u'}{\partial\xi}\right]=0, \tag{2}$$

$$\frac{h^2}{3}\nabla_1^4 w'+(1-\sigma^2)\frac{S}{Eh}\frac{\partial^2 w'}{\partial\xi\,\partial\eta}-\frac{1}{\rho}\left[\left(\frac{\partial v'}{\partial\eta}-\frac{w'}{\rho}\right)+\sigma\frac{\partial u'}{\partial\xi}\right]=0. \tag{3}$$

We will now put the equations (1), (2) and (3) into non-dimensional form by means of the substitutions

$$\xi=\frac{bx}{\pi},\quad \eta=\frac{by}{\pi}-\frac{b}{2}.$$

We get

$$\frac{\partial}{\partial x}\left[\frac{\partial u'}{\partial x}+\sigma\left(\frac{\partial v'}{\partial y}-\frac{bw'}{\pi\rho}\right)\right]+\frac{(1-\sigma)}{2}\frac{\partial}{\partial y}\left(\frac{\partial v'}{\partial x}+\frac{\partial u'}{\partial y}\right)=0, \tag{4}$$

$$\frac{(1-\sigma)}{2}\frac{\partial}{\partial x}\left(\frac{\partial v'}{\partial x}+\frac{\partial u'}{\partial y}\right)+\frac{\partial}{\partial y}\left[\left(\frac{\partial v'}{\partial y}-\frac{bw'}{\pi\rho}\right)+\sigma\frac{\partial u'}{\partial x}\right]=0, \tag{5}$$

$$\frac{h^2}{3}\nabla_1^4 w'+(1-\sigma^2)\left(\frac{b}{\pi}\right)^2\frac{S}{Eh}\frac{\partial^2 w'}{\partial x\,\partial y}-\frac{b^3}{\pi^3\rho}\left[\left(\frac{\partial v'}{\partial y}-\frac{bw'}{\pi\rho}\right)+\sigma\frac{\partial u'}{\partial x}\right]=0. \tag{6}$$

We now effect a simplification by introducing a function ψ (cf. Dean 1926; Timoshenko 1934, chap. 1) defined by the equation

$$\frac{(1-\sigma)}{2}\left(\frac{\partial u'}{\partial y}+\frac{\partial v'}{\partial x}\right)=-\pi\frac{\partial^2\psi}{\partial x\,\partial y}. \tag{7}$$

From (4), (5), (6) and (7) we deduce the following four equations:

$$\frac{\partial u'}{\partial x}+\sigma\left(\frac{\partial v'}{\partial y}-\frac{bw'}{\pi\rho}\right)=\pi\frac{\partial^2\psi}{\partial y^2}, \tag{8}$$

$$\left(\frac{\partial v'}{\partial y}-\frac{bw'}{\pi\rho}\right)+\sigma\frac{\partial u'}{\partial x}=\pi\frac{\partial^2\psi}{\partial x^2}, \tag{9}$$

$$\nabla_1^4\psi+R\frac{\partial^2 w'}{\partial x^2}=0, \tag{10}$$

$$\nabla_1^4 w' - P\frac{\partial^2 \psi}{\partial x^2} + Q\frac{\partial^2 w'}{\partial x \partial y} = 0, \tag{11}$$

where
$$P = \frac{3b}{\rho h^2}\left(\frac{b}{\pi}\right)^2, \quad Q = \frac{2S}{D}\left(\frac{b}{\pi}\right)^2, \quad R = \frac{(1-\sigma^2)}{b\rho}\left(\frac{b}{\pi}\right)^2. \tag{12}$$

4—Case 1—*Edges Fixed and Simply Supported*—We shall now find a solution of the above equations which satisfies the boundary conditions*

$$\left.\begin{aligned} u' = v' &= 0, \\ w' = \frac{\partial^2 w'}{\partial y^2} &= 0, \end{aligned}\right\} \text{ for } y = 0,\ \pi. \tag{13}$$

Since we are omitting the corresponding conditions over the two curved ends, it is possible to express w' and ψ in the form

$$w' = w_1' \cos mx + w_2' \sin mx,$$
$$\psi = \psi_1 \cos mx + \psi_2 \sin mx,$$

whence introducing the two complex quantities W' and Ψ defined by the relations

$$W' = w_1' + iw_2', \quad \Psi = \psi_1 + i\psi_2.$$

We are enabled to express w' and ψ more compactly as

$$w' = (R)W'e^{-imx}, \quad \psi = (R)\Psi e^{-imx}. \tag{14}$$

In (14) W' and Ψ are functions of y only, and m is real but otherwise unspecified.

Substituting for w' and ψ in (10) and (11), we derive the two fourth order differential equations

$$\left(\frac{d^2}{dy^2} - m^2\right)^2 \Psi - Rm^2 W' = 0,$$

$$\left(\frac{d^2}{dy^2} - m^2\right)^2 W' + Pm^2\Psi - imQ\frac{dW'}{dy} = 0.$$

Since they are linear and have constant coefficients, their general solution is of the form

$$\Psi = \sum_{r=1}^{\infty} a_r e^{\tau_r y}, \quad W' = \sum_{r=1}^{\infty} b_r e^{\tau_r y},$$

where the a_r's and b_r's are interdependent complex constants and the τ_r's are the roots of the algebraic equation

$$(\tau^2 - m^2)^2\{(\tau^2 - m^2)^2 - imQ\tau\} + PRm^4 = 0.$$

---

* The expression for the couple in general contains two terms, but since w' is zero over the two straight edges the second term vanishes.

From this equation for τ it can be seen that the algebra involved by this method of solution would be so heavy as to make the method scarcely practicable, but before considering any alternative we make an important observation which we shall use later, namely, that the general solution for Ψ and W' is such that Ψ and W' and all their derivatives are continuous functions of y in (0π).

We shall now proceed to develop a method of solving the equations (10) and (11) entirely different from that just outlined above. From the last two boundary conditions in (13), we deduce

$$w_s' = \frac{d^2 w_s'}{dy^2} = 0, \quad [s = 1, 2] \text{ for } y = 0, \pi. \tag{15}$$

Now consider the fourth derivative of w_s' ($s = 1$ or 2) with respect to y in $<0\pi>$. It is continuous in that range, and hence can be expanded as a Fourier series of sines throughout (0π) in the form

$$\frac{d^4 w_s'}{dy^4} = \sum_{r=1}^{\infty} r^4 A_{sr} \sin ry, \tag{16}$$

where by a well-known property of Fourier* series $r^4 \mid A_{sr} \mid < K/r$, K being some positive number independent of r. Making use of the above inequality and another property of Fourier series, we integrate (16) twice, and deduce that in (0π)

$$\frac{d^2 w_s'}{dy^2} = -\sum_{r=1}^{\infty} r^2 A_{sr} \sin ry + B + Cy. \tag{17}$$

Since each side of (16) is a continuous function of y, it follows that the equation is true, not only in the open interval (0π) but in the closed interval $<0\pi>$ as well. We now put y equal to 0, π, whence it follows from (15) that B and C are zero. Integrating (17) twice and adopting the same reasoning as above, we deduce that

$$w_s' = \sum_{r=1}^{\infty} A_{\;r} \sin ry \text{ in } <0\pi>.$$

Reversing this process we obtain the result we want, namely, that we can assume a Fourier series in sines for w_s' in $<0\pi>$ which can be differentiated four times term by term.

Introducing A_r defined by the relation

$$A_r = A_{1r} + A_{2r},$$

---

* For the properties of Fourier series used in this paper see Carslaw (1930, chaps. 7, 8).

and substituting for w' in (10), we have the following equation from which to determine Ψ:

$$\frac{d^4\Psi}{dy^4}-2m^2\frac{d^2\Psi}{dy^2}+m^4\Psi = m^2R\sum_{r=1}^{\infty}A_r\sin ry.$$

The general solution is

$$\Psi = \{A\cosh my+B\sinh my\}+y\{C\cosh my+D\sinh my\} \\ +m^2R\sum_{r=1}^{\infty}\frac{A_r\sin ry}{(r^2+m^2)^2}. \tag{18}$$

We will now obtain A, B, C, D in terms of A_r by virtue of the first two boundary conditions in (13). Since these involve u' and v' it is necessary to solve for them from the equations (7), (8), (9). Assuming a form of solution given by

$$\left.\begin{aligned} u' &= u_1'\cos mx+u_2'\sin mx, \\ v' &= v_1'\cos mx+v_2'\sin mx, \end{aligned}\right\} \tag{19}$$

where u_1', u_2', v_1', v_2' are independent of x; and introducing U' and V' defined by

$$U' = u_1'+iu_2', \quad V' = v_1'+iv_2',$$

so that $$u' = (R)\,U'e^{-imx}, \quad v' = (R)\,V'e^{-imx},$$

we obtain from (7), (8), (9), (19) the three following equations:*

$$\frac{(1-\sigma)}{2}\left\{\frac{dU'}{dy}-imV'\right\} = i\pi m\frac{d\Psi}{dy}, \tag{20}$$

$$-imU'+\sigma\frac{dV'}{dy} = \pi\frac{d^2\Psi}{dy^2}+\frac{\sigma b}{\pi\rho}W', \tag{21}$$

$$\frac{dV'}{dy}-i\sigma mU' = -\pi m^2\Psi+\frac{bW'}{\pi\rho}. \tag{22}$$

Manipulating these equations we have

$$mU'(1-\sigma^2) = i\pi\frac{d^2\Psi}{dy^2}+i\pi\sigma m^2\Psi, \tag{23}$$

$$mV' = -i\frac{dU'}{dy}-\frac{2\pi m}{(1-\sigma)}\frac{d\Psi}{dy}. \tag{24}$$

---

* At first sight it might be thought that we had found three equations to be satisfied by the two unknown functions U' and V', but it must be remembered that of the five equations (7) to (11) only four are independent. Since we are treating (10) and (11) as independent, it follows that of the three equations (7), (8), (9), and therefore of the three equations (20), (21), (22), only two are independent.

The first two boundary conditions in (13) can be expressed in the form

$$U' = V' = 0, \text{ for } y = 0, \pi,$$

whence from (23) and (24) we deduce

$$\left.\begin{aligned} \frac{d^2\Psi}{dy^2} + \sigma m^2 \Psi &= 0, \\ \frac{d^3\Psi}{dy^3} - (2+\sigma)\, m^2 \frac{d\Psi}{dy} &= 0, \end{aligned}\right\} \text{ for } y = 0, \pi. \tag{25}$$

Substituting for Ψ given by (18) in the equation (25), we have the following four equations from which to determine A, B, C, D:

$$Am(1+\sigma) + 2D = 0, \tag{26}$$

$$\begin{aligned} Am(1+\sigma)\cosh m\pi + Bm(1+\sigma)\sinh m\pi \\ + C\{2\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\} \\ + D\{2\cosh m\pi + m\pi(1+\sigma)\sinh m\pi\} = 0, \end{aligned} \tag{27}$$

$$-Bm(1+\sigma) + C(1-\sigma) = RX, \tag{28}$$

$$\begin{aligned} -Am(1+\sigma)\sinh m\pi - Bm(1+\sigma)\cosh m\pi \\ + C\{(1-\sigma)\cosh m\pi - m\pi(1+\sigma)\sinh m\pi\} \\ + D\{(1-\sigma)\sinh m\pi - m\pi(1+\sigma)\cosh m\pi\} = RY, \end{aligned} \tag{29}$$

where

$$\left.\begin{aligned} X &= \sum_{r=1}^{\infty} \frac{A_r\{r^3 + (2+\sigma)\, m^2 r\}}{(r^2+m^2)^2} = \sum_{r=1}^{\infty} A_r \alpha_r, \\ Y &= \sum_{r=1}^{\infty} (-)^r \frac{A_r\{r^3 + (2+\sigma) m^2 r\}}{(r^2+m^2)^2} = \sum_{r=1}^{\infty} (-)^r A_r \alpha_r. \end{aligned}\right\} \tag{30}$$

The second half of the two equations (30) is to be regarded as defining α_r which it will be convenient to use later.

Solving (26) to (29) we deduce

$$A = \frac{A'}{\Delta}, \quad B = \frac{B'}{\Delta}, \quad C = \frac{C'}{\Delta}, \quad D = \frac{D'}{\Delta}, \tag{31}$$

where

$$\Delta = m^2(1+\sigma)^2\{(3-\sigma)^2 \sinh^2 m\pi - m\pi^2(1+\sigma)^2\}, \tag{32}$$

$$\begin{aligned} A' = 2RXm(1+\sigma)\{(3-\sigma)\sinh m\pi \cosh m\pi + m\pi(1+\sigma)\} \\ - 2RYm(1+\sigma)\{(3-\sigma)\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\}, \end{aligned} \tag{33}$$

$$\begin{aligned} B' = RXm(1+\sigma)\{m^2\pi^2(1+\sigma)^2 - 2(3-\sigma)\sinh^2 m\pi\} \\ - RYm^2\pi(1+\sigma)^2(1-\sigma)\sinh m\pi, \end{aligned} \tag{34}$$

$$C' = RXm^2(1+\sigma)^2(3-\sigma)\sinh^2 m\pi - RYm^3\pi(1+\sigma)^3 \sinh m\pi, \tag{35}$$

$$\begin{aligned} D' = -RXm^2(1+\sigma)^2\{(3-\sigma)\sinh m\pi \cosh m\pi + m\pi(1+\sigma)\} \\ + RYm^2(1+\sigma^2)\{(3-\sigma)\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\}. \end{aligned} \tag{36}$$

After substituting the values of A, B, C, D, obtained from (31) to (36) in the equation (18), the resulting expressions for w' and ψ given by (14) will satisfy all the fundamental equations and boundary conditions except the equation (11). We now therefore substitute for w' and ψ in (11) and equate to zero the coefficient of e^{-imx}. We get

$$\sum_{r=1}^{\infty} A_r(r^2+m^2)^2 \sin ry - imQ \sum_{r=1}^{\infty} rA_r \cos ry + PRm^4 \sum_{r=1}^{\infty} \frac{A_r \sin ry}{(r^2+m^2)^2} + Pm^2\{A\cosh my + B\sinh my + y(C\cosh my + D\sinh my)\} = 0, \quad (37)$$

which must be satisfied for all y in (0π).

Now before we can proceed with the equation (37), it is necessary to express all the terms in y as Fourier series of sines.

Expressing $\cos ny$ as a sine series of the form

$$\cos ny = \sum_{r=1}^{\infty} \beta_r \sin ry, \quad [0 < y < \pi] \quad (38)$$

we deduce that the Fourier constants β_r are given by

$$\beta_r = \frac{1}{\pi}\left[\frac{1-\cos(r+n)\pi}{r+n} + \frac{1-\cos(r-n)\pi}{r-n}\right]. \quad (39)$$

Hence from (38) and (39) the coefficient of $\sin ry$ in the expansion of $\sum_{r=1}^{\infty} rA_r \cos ry$ is

$$\frac{1}{\pi}\sum_{n=1}^{\infty} nA_n\left[\frac{1-\cos(r+n)\pi}{r+n} + \frac{1-\cos(r-n)\pi}{r-n}\right]. \quad (40)$$

By straightforward evaluation it follows that the coefficient of A_n in (40), which we shall denote by c_{rn}, is given by the adjoining scheme:

	A_1	A_2	A_3	A_4
$\sin y$	0	$-\frac{8}{3\pi}$	0	$-\frac{16}{15\pi}$
$\sin 2y$	$\frac{8}{3\pi}$	0	$-\frac{24}{5\pi}$	0
$\sin 3y$	0	$\frac{24}{5\pi}$	0	$-\frac{48}{7\pi}$
$\sin 4y$	$\frac{16}{15\pi}$	0	$\frac{48}{7\pi}$	0

(41)

It now remains to expand in a sine series in (0π) the expression

$$(A\cosh my + B\sinh my) + y(C\cosh my + D\sinh my).$$

Writing it as $\sum_{r=1}^{\infty}\kappa_r \sin ry$ we get

$$\kappa_r = \frac{2}{\pi}\int_0^{\pi} \sin ry\,[A\cosh my + B\sinh my + y\,(C\cosh my + D\sinh my)]\,dy. \tag{42}$$

Denoting the coefficients of A, B, C, D in (42) by I_{1r}, I_{2r}, I_{3r}, I_{4r}, we have

$$\frac{\pi}{2}I_{1r} = \frac{r}{(r^2+m^2)}[1-\cosh m\pi\cos r\pi], \tag{43}$$

$$\frac{\pi}{2}I_{2r} = -\frac{r}{(r^2+m^2)}\sinh m\pi\cos r\pi, \tag{44}$$

$$\frac{\pi}{2}I_{3r} = -\frac{r\cos r\pi}{(r^2+m^2)^2}[\pi(r^2+m^2)\cosh m\pi - 2m\sinh m\pi], \tag{45}$$

$$\frac{\pi}{2}I_{4r} = -\frac{r}{(r^2+m^2)^2}[\pi(r^2+m^2)\cos r\pi\sinh m\pi - 2m(\cos r\pi\,\cosh m\pi - 1)]. \tag{46}$$

It is now essential to get κ_r expressed in as simple a form as possible, and for this purpose we introduce L_r, M_r, defined by the equation

$$AI_{1r} + BI_{2r} + CI_{3r} + DI_{4r} = \frac{R}{m^2}(L_r X + M_r Y), \tag{47}$$

so that we have

$$\kappa_r = \frac{R}{m^2}\left[L_r\sum_{n=1}^{\infty}A_n\alpha_n + M_r\sum_{n=1}^{\infty}(-)^n A_n\alpha_n\right]. \tag{48}$$

From (31) to (36) and (43) to (47), we deduce by straightforward algebra that

$$L_r = \frac{4m\alpha_r}{\pi(1+\sigma)}\,\frac{\left[\begin{array}{l}\{(3-\sigma)\cosh m\pi\,\sinh m\pi + m\pi\,(1+\sigma)\}\\ \qquad\qquad -\cos r\pi\{(3-\sigma)\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\}\end{array}\right]}{(3-\sigma)^2\sinh^2 m\pi - m^2\pi^2(1+\sigma)^2}, \tag{49}$$

$$M_r = \frac{4m\alpha_r}{\pi(1+\sigma)}\,\frac{\left[\begin{array}{l}\cos r\pi\{(3-\sigma)\cosh m\pi\sinh m\pi + m\pi(1+\sigma)\}\\ \qquad\qquad -\{(3-\sigma)\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\}\end{array}\right]}{(3-\sigma)^2\sinh^2 m\pi - m^2\pi^2(1+\sigma)^2}, \tag{50}$$

and from (49) and (50) it follows that

$$L_r = (-)^r M_r. \tag{51}$$

On making use of equations (40) and (48), (37) can be expressed in terms of sines of multiples of y only, and since the equation is true for all y in (0π), the coefficient of $\sin ry (r = 1, 2, \ldots, \infty)$ must vanish. This gives us an infinite system of equations, linear in the A's, namely,

$$(r^2+m^2)^2 A_r - imQ \sum_{n=1}^{\infty} c_{rn} A_n + \frac{PRm^4}{(r^2+m^2)^2} A_r + PR\left[L_r \sum_{n=1}^{\infty} \alpha_n A_n + M_r \sum_{n=1}^{\infty} (-)^n \alpha_n A_n\right] = 0. \quad (52)$$

In general the only solution of the above equations is that in which all the A's are zero. If however the infinite determinant formed by eliminating the A's vanishes, this is no longer the case. Since on physical grounds a solution of this kind is only possible when the plate is in limiting equilibrium, it follows that by equating the infinite determinant to zero, we possess an equation from which to determine the critical value of S. This equation, being of indefinitely large order, will have an infinite number of roots, but since for practical purposes we are only interested in the shearing force necessary to produce the *most favoured* type of distortion, it is only with the smallest of these roots that we are concerned.

We now make the substitution

$$(r^2+m^2)^2 + \frac{PRm^4}{(r^2+m^2)^2} + PR\alpha_r(L_r + (-)^r M_r) = Z_r, \quad (53)$$

whence the required equation is

$$\begin{vmatrix} Z_1 & -imQc_{12} & PR\alpha_3(L_1 - M_1) & -imQc_{14} \\ -imQc_{21} & Z_2 & -imQc_{23} & PR\alpha_4(L_2 + M_2) \\ PR\alpha_1(L_3 - M_3) & -imQc_{32} & Z_3 & -imQC_{34} \\ -imQc_{41} & PR\alpha_2(L_4 + M_4) & -imQc_{43} & Z_4 \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0, \quad (54)$$

after making use of equation (51), and the fact that half the c's vanish. Equation (54) is only formal, however, and it is first necessary to prove that the determinant is convergent before we are at liberty to start evaluating S.

Divide the rth row and column by r^2. Then the product of the diagonal elements converges absolutely, and it remains to show that the sum of the non-diagonal elements also converges absolutely (Whittaker and Watson 1927, § 2·81).

Since from (30), (49), (50), α_r, L_r, M_r, are $O(1/r)$, the sum of the modulus of the terms involving α, L, M is less than

$$K\left\{\sum_{r=1}^{\infty} \frac{1}{r^2}\right\}^2 \text{ (K being some constant)} \quad (55)$$

and therefore convergent; but the terms containing the c's require a more detailed consideration. The sum of the modulus of these terms is

$$Qm \sum_{r=1}^{\infty} \sum_{n=1}^{\infty} \frac{|c_{rn}|}{r^2 n^2},$$

which, on making use of (41), is less than

$$\frac{2Qm}{\pi}\left\{\sum_{r=1}^{\infty}\frac{1}{r^2}\right\}^2 + \frac{2Qm}{\pi}\sum_{r=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{r^2 n\,|r-n|}. \quad (r \neq n) \tag{56}$$

Now consider $\sum_{n=1}^{\infty} \frac{1}{n\,|r-n|}$ $(r \neq n)$. It is equal to

$$\sum_{n=1}^{r-1}\frac{1}{n(r-n)} + \sum_{n=r+1}^{\infty}\frac{1}{n(n-r)}. \tag{57}$$

By the Cauchy-Maclaurin integral test, we can prove at once that the expression (57) possesses an upper bound independent of r; hence the double series in (56) is convergent, and hence from (55) and (56) the sum of the non-diagonal elements is absolutely convergent—which is what we wanted to prove.

Since we cannot obtain the roots of (54) directly, we approximate to them by considering the values of S obtained by replacing the left-hand side of (54) by the finite determinant which contains successively the first 4, 9, 16, ... elements of the determinant in (54).

We will now proceed to evaluate these determinants, considering the accuracy of the results obtained in this way after we have actually found them.

5—As a *first* approximation, we consider the determinant formed by taking the first four elements of the determinant in (54). We have

$$\begin{vmatrix} Z_1 & -imQc_{12} \\ -imQc_{21} & Z_2 \end{vmatrix} = 0,$$

giving on evaluation

$$\frac{8Qm}{3\pi} = (Z_1 Z_2)^{\frac{1}{2}}.$$

Substituting for Q from (12), we deduce finally

$$S = \frac{3D\pi^3}{16mb^2}(Z_1 Z_2)^{\frac{1}{2}},$$

where Z_1, Z_2 are given by (53), and m has that particular value which makes S a minimum.

For a *second* approximation we take the first nine elements in (54) giving ultimately, after making use of (41) and (51),

$$S = \frac{D\pi^2}{2mb^2}\left\{\frac{Z_1 Z_2 Z_3 - 4Z_2\alpha_3^2 L_1^2 P^2 R^2}{c_{23}^2 Z_1 + c_{12}^2 Z_3 + 4c_{12}c_{23}\alpha_3 L_1 PR}\right\}^{\frac{1}{2}}. \tag{58}$$

Proceeding in a similar manner, we obtain from the fourth order determinant appropriate to the *third* approximation a quadratic for Q^2. This we solve for the smaller of the two roots, but as the coefficients in the quadratic are long and complicated, the process is not set down here.

The limiting value of S, as calculated from the third approximation when ρ is infinite,* is in very close agreement with the value obtained by Southwell and Skan (1924) in their paper. Calling this value S_0, k, in Table I, gives the ratio of the shearing force, calculated for any particular case, to S_0.

TABLE I

	First approximation		Second approximation		Third approximation	
$\sqrt{RP}$	m^2	k	m^2	k	m^2	k
0	0·59	1·046	0·63	1·000	0·63	1·000
1	0·77	1·199	0·83	1·118	0·83	1·116
2	1·05	1·523	1·15	1·342	—	—
3	1·35	1·914	1·35	1·574	—	—
4	1·6	2·360	1·55	1·802	1·6	1·785
5·5	1·9	3·128	1·75	2·133	—	—
7	2·3	4·017	1·75	2·458	1·85	2·348

We do not consider the significance of the above results† until we have solved the problem for an alternative set of boundary conditions. The results are shown graphically in figs. 2, 3.

6—Case II—*Edges Fixed and Clamped*—In this event the boundary conditions to be satisfied are

$$\left.\begin{aligned} u' = v' &= 0, \\ w' = \frac{\partial w'}{\partial y} &= 0, \end{aligned}\right\} \text{for } y = 0, \pi. \tag{59}$$

The method of solution adopted in this case is similar to that explained in § 4, and as in that case it is necessary to start by considering to what extent we can assume a differentiable Fourier series for w'.

---

* $52{\cdot}85\ D/b^2$.

† The value of σ is taken to be 0·25.

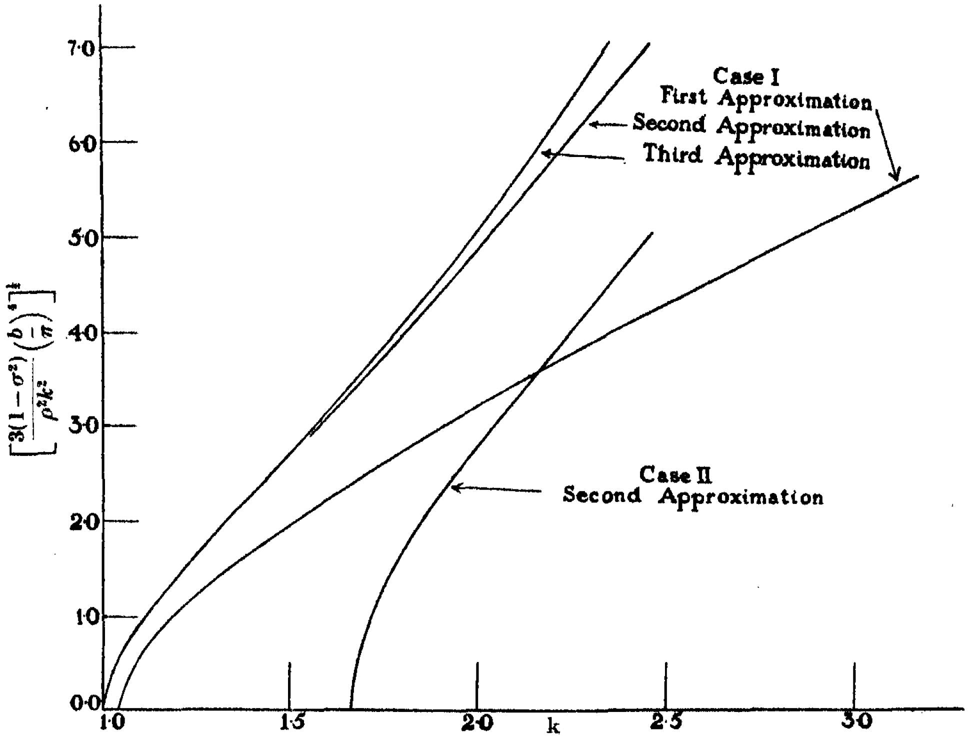

Case I
First Approximation
Second Approximation
Third Approximation
Case II
Second Approximation
$\left[\frac{3(1-\sigma^2)}{\rho^2 k^2}\left(\frac{b}{\pi}\right)^4\right]^{\frac{1}{2}}$
7·0
6·0
5·0
4·0
3·0
2·0
1·0
0·0
1·0
1·5
2·0
k
2·5
3·0

FIG. 2

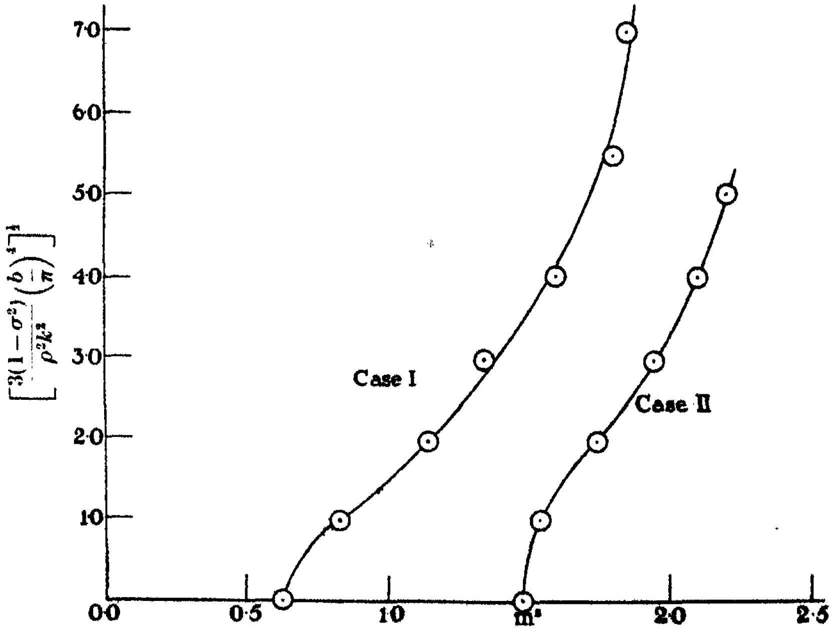

Case I
Case II
$\left[\frac{3(1-\sigma^2)}{\rho^2 k^2}\left(\frac{b}{\pi}\right)^4\right]^{\frac{1}{2}}$
7·0
6·0
5·0
4·0
3·0
2·0
1·0
0·0
0·5
1·0
m²
2·0
2·5

FIG. 3

We express w' and ψ in the form assumed at the beginning of Case I, but the last two boundary conditions in (59) now give rise to

$$w'_s = \frac{dw'_s}{dy} = 0, \quad [s = 1, 2] \text{ for } y = 0, \pi. \tag{60}$$

Expanding the fourth derivative of w'_s as a sine series in (0π), we have

$$\frac{d^4 w'_s}{dy^4} = \sum_{r=1}^{\infty} r^4 A_{sr} \sin ry, \tag{61}$$

and A_{sr} satisfies the inequality

$$r^4 | A_{sr} | < \frac{K}{r}.$$

Integrating equation (61) three and four times, we deduce for $0 \leqslant y \leqslant \pi$

$$\left.\begin{aligned} \frac{dw'_s}{dy} &= \sum_{r=1}^{\infty} rA_{sr} \cos ry + E_s + 2F_s y + 3G_s y^2, \\ w'_s &= \sum_{r=1}^{\infty} A_{sr} \sin ry + H_s + E_s y + F_s y^2 + G_s y^3, \end{aligned}\right\} \tag{62}$$

where H_s, E_s, F_s, G_s are constants. From (60) these are given by

$$H_s = 0, \qquad E_s + \pi F_s + \pi^2 G_s = 0,$$

$$E_s = -\sum_{r=1}^{\infty} rA_{sr}, \quad E_s + 2\pi F_s + 3\pi^2 G_s = -\sum_{r=1}^{\infty} r \cos r\pi A_{sr},$$

leading to

$$\left.\begin{aligned} H_s &= 0, & F_s &= \frac{1}{\pi} \sum_{r=1}^{\infty} \{2 + (-)^r\} rA_{sr}, \\ E_s &= -\sum_{r=1}^{\infty} rA_{sr}, & G_s &= -\frac{1}{\pi^2} \sum_{r=1}^{\infty} \{1 + (-)^r\} rA_{sr}. \end{aligned}\right\} \tag{63}$$

Hence we can assume for w'_s a form given by (62) and (63), where the sine series can be differentiated term by term four times. Introducing W', Ψ and A_r, as defined in Case I, w' and ψ can still be expressed in the form given by (14). We now introduce the additional complex quantities E, F and G, defined by the relations

$$E = E_1 + iE_2, \quad F = F_1 + iF_2, \quad G = G_1 + iG_2.$$

Substituting for w' in (10), we have the following differential equation for Ψ:

$$\frac{d^4\Psi}{dy^4} - 2m^2 \frac{d^2\Psi}{dy^2} + m^4\Psi = m^2 R\left\{\sum_{r=1}^{\infty} A_r \sin ry + Ey + Fy^2 + Gy^3\right\},$$

which has for its general solution

$$\Psi = \{A\cosh my + B\sinh my\} + y\{C\cosh my + D\sinh my\} + m^2R\left[\sum_{r=1}^{\infty}\frac{A_r\sin ry}{(r^2+m^2)^2} + \frac{1}{m^4}\left\{\frac{4F}{m^2} + y\left(E + \frac{12G}{m^2}\right) + Fy^2 + Gy^3\right\}\right]. \qquad (64)$$

It is now necessary to obtain the constants A, B, C, D from the first two boundary conditions in (59). The first part of this process is identical with that contained in equations (19) to (25) in Case I and so is not repeated here, but on substituting for Ψ given by (64) in (25), we deduce four equations from which to determine A, B, C, D, which are slightly different from the corresponding equations (26) to (29). The left-hand sides of the two sets of equations are the same, but in place of 0, 0, RX, RY on the right-hand side of (26), (27), (28), (29), we now have $R\theta$, $R\phi$, $R(\chi+\psi)$, $R(\gamma+\Omega)$ for the corresponding equations in Case II, where

$$\psi = -\frac{(2+\sigma)}{m^2}\sum_{r=1}^{\infty} rA_r - \frac{6(3+2\sigma)}{\pi^2 m^4}\sum_{r=1}^{\infty}\{1+(-)^r\}rA_r = \sum_{r=1}^{\infty} A_r\gamma_r,$$

$$\Omega = -\frac{(2+\sigma)}{m^2}\sum_{r=1}^{\infty}(-)^r rA_r - \frac{6(3+2\sigma)}{\pi^2 m^4}\sum_{r=1}^{\infty}\{1+(-)^r\}rA_r = \sum_{r=1}^{\infty}(-)^r A_r\gamma_r,$$

$$\theta = -\frac{2(1+2\sigma)}{\pi m^3}\left\{\sum_{r=1}^{\infty}[2+(-)^r]rA_r\right\} = \sum_{r=1}^{\infty} A_r\mu_r,$$

$$\phi = \frac{2(1+2\sigma)}{\pi m^3}\left\{\sum_{r=1}^{\infty}[1+2(-)^r]rA_r\right\} = -\sum_{r=1}^{\infty}(-)^r A_r\mu_r.$$

The two new quantities introduced, γ_r and μ_r, are to be regarded as defined by the second half of the above equations.

Solving the four equations analogous to (26) to (29), A, B, C, D are given by (31), (32), and

$$\begin{aligned} A' = {} & 2R(X+\psi)m(1+\sigma)\{(3-\sigma)\sinh m\pi\cosh m\pi + m\pi(1+\sigma)\} \\ & - 2R(Y+\Omega)m(1+\sigma)\{(3-\sigma)\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\} \\ & + R\theta m(1+\sigma)\{(3-\sigma)(1-\sigma)\sinh^2 m\pi - m^2\pi^2(1+\sigma)^2\} \\ & - 2R\phi m^2\pi(1+\sigma)^2\sinh m\pi, \end{aligned} \qquad (65)$$

$$\begin{aligned} B' = {} & R(X+\psi)m(1+\sigma)\{m^2\pi^2(1+\sigma)^2 - 2(3-\sigma)\sinh^2 m\pi\} \\ & - R(Y+\Omega)m^2\pi(1+\sigma)^2(1-\sigma)\sinh m\pi \\ & + R\theta m(1-\sigma^2)\{m\pi(1+\sigma) - (3-\sigma)\sinh m\pi\cosh m\pi\} \\ & + R\phi m(1-\sigma^2)\{(3-\sigma)\sinh m\pi - m\pi(1+\sigma)\cosh m\pi\}, \end{aligned} \qquad (66)$$

$$\begin{aligned} C' = {} & R(X+\psi)\,m^2(1+\sigma)^2(3-\sigma)\sinh^2 m\pi \\ & -R(Y+\Omega)\,m^3\pi(1+\sigma)^3\sinh m\pi \\ & +R\theta m^2(1+\sigma)^2\{m\pi(1+\sigma)-(3-\sigma)\sinh m\pi\cosh m\pi\} \\ & +R\phi m^2(1+\sigma)^2\{(3-\sigma)\sinh m\pi - m\pi(1+\sigma)\cosh m\pi\}, \end{aligned} \tag{67}$$

$$\begin{aligned} D' = {} & -R(X+\psi)\,m^2(1+\sigma)^2\{(3-\sigma)\sinh m\pi\cosh m\pi + m\pi(1+\sigma)\} \\ & +R(Y+\Omega)\,m^2(1+\sigma)^2\{(3-\sigma)\sinh m\pi + m\pi(1+\sigma)\cosh m\pi\} \\ & +R\theta m^2(1+\sigma)^2(3-\sigma)\sinh^2 m\pi + R\phi m^3\pi(1+\sigma)^3\sinh m\pi. \end{aligned} \tag{68}$$

Since the expressions obtained for w' and ψ now satisfy all the required conditions except equation (11), it remains to substitute for w' and ψ in that equation and equate to zero the coefficient of e^{-imx}. Corresponding to (37), we get for all y in (0π)

$$\begin{aligned} & \sum_{r=1}^{\infty} A_r(r^2+m^2)^2\sin ry + \{m^4(Ey+Fy^2+Gy^3)-2m^2(2F+6Gy)\} \\ & \quad -imQ\sum_{r=1}^{\infty} rA_r\cos ry - imQ(E+2Fy+3Gy^2) + PRm^4\sum_{r=1}^{\infty} A_r\frac{\sin ry}{(r^2+m^2)^2} \\ & \quad +Pm^2\{A\cosh my + B\sinh my + y(C\cosh my + D\sinh my)\} \\ & \quad +PR\left\{\frac{4F}{m^2}+y\left(E+\frac{12G}{m^2}\right)+Fy^2+Gy^3\right\} = 0. \end{aligned} \tag{69}$$

It is now necessary to expand the left-hand side of equation (69) in terms of sines of multiples of y. For the cosine series, and the terms containing A, B, C, D, we make use of the work done in Case I and expressed in equations (38) to (46); but for the terms involving E, F, G we have to construct schemes analogous to (41).

Expressing 1, y, y^2, y^3 as Fourier series of sines in (0π), we deduce

$$\left.\begin{aligned} 1 &= \frac{2}{\pi}\sum_{r=1}^{\infty}\frac{\{1-(-)^r\}}{r}\sin ry, \\ y &= -2\sum_{r=1}^{\infty}\frac{(-)^r}{r}\sin ry, \\ y^2 &= -2\sum_{r=1}^{\infty}\left[(-)^r\frac{\pi}{r}+\frac{2\{1-(-)^r\}}{\pi r^3}\right]\sin ry, \\ y^3 &= -2\sum_{r=1}^{\infty}(-)^r\left\{\frac{\pi^2}{r}-\frac{6}{r^3}\right\}\sin ry. \end{aligned}\right\} \tag{70}$$

Hence from (63) and (70), the coefficients of $\sin ry$ in the expansion of $Ey+Fy^2+Gy^3$, $F+3Gy$, $E+2Fy+3Gy^2$ are respectively

$$-\frac{4}{\pi^2 r^3}\{1-(-)^r\}\sum_{n=1}^{\infty}\{2+(-)^n\}nA_n - \frac{(-)^r\,12}{\pi^2 r^3}\sum_{n=1}^{\infty}\{1+(-)^n\}nA_n, \tag{71}$$

$$\frac{2}{\pi^2 r}\{1-(-)^r\}\sum_{n=1}^{\infty}\{2+(-)^n\}nA_n+\frac{(-)^r 6}{\pi^2 r}\sum_{n=1}^{\infty}\{1+(-)^r\}nA_n, \tag{72}$$

$$-\frac{2}{\pi r}\sum_{n=1}^{\infty}nA_n+\frac{(-)^r 2}{\pi r}\sum_{n=1}^{\infty}(-)^n nA_n+\frac{12}{\pi^3 r^3}\{1-(-)^r\}\sum_{n=1}^{\infty}\{1+(-)^n\}nA_n. \tag{73}$$

Denoting the coefficients of A_n in (71), (72), (73), by e_{rn}, f_{rn}, g_{rn}, we derive the following schemes for these three sets of quantities:

	A_1	A_2	A_3	A_4
$\sin y$	$-\frac{8}{\pi^2}$	0	$-\frac{24}{\pi^2}$	0
$\sin 2y$	0	$-\frac{6}{\pi^2}$	0	$-\frac{12}{\pi^2}$
$\sin 3y$	$-\frac{8}{27\pi^2}$	0	$-\frac{8}{9\pi^2}$	0
$\sin 4y$	0	$-\frac{3}{4\pi^2}$	0	$-\frac{3}{2\pi^2}$

(74)

	A_1	A_2	A_3	A_4
$\sin y$	$\frac{4}{\pi^2}$	0	$\frac{12}{\pi^2}$	0
$\sin 2y$	0	$\frac{12}{\pi^2}$	0	$\frac{24}{\pi^2}$
$\sin 3y$	$\frac{4}{3\pi^2}$	0	$\frac{4}{\pi^2}$	0
$\sin 4y$	0	$\frac{6}{\pi^2}$	0	$\frac{12}{\pi^2}$

(75)

	A_1	A_2	A_3	A_4
$\sin y$	0	$-\frac{8}{\pi}+\frac{96}{\pi^3}$	0	$-\frac{16}{\pi}+\frac{192}{\pi^3}$
$\sin 2y$	$-\frac{2}{\pi}$	0	$-\frac{6}{\pi}$	0
$\sin 3y$	0	$-\frac{8}{3\pi}+\frac{32}{9\pi^3}$	0	$-\frac{16}{3\pi}+\frac{64}{9\pi^3}$
$\sin 4y$	$-\frac{1}{\pi}$	0	$-\frac{3}{\pi}$	0

(76)

Following the development of Case I, it is necessary to simplify κ_r, and we accordingly proceed as in that case. Two new quantities U_r, V_r are introduced, being defined by the equation

$$AI_{1r}+BI_{2r}+CI_{3r}+DI_{4r}=\frac{R}{m^2}[L_r(X+\psi)+M_r(Y+\Omega)+U_r\theta+V_r\phi],$$

whence
$$\kappa_r=\frac{R}{m^2}\left[L_r\left\{\sum_{n=1}^{\infty}A_n(\alpha_n+\gamma_n)\right\}+M_r\left\{\sum_{n=1}^{\infty}(-)^nA_n(\alpha_n+\gamma_n)\right\}\right.$$
$$\left.+U_r\sum_{n=1}^{\infty}A_n\mu_n-V_r\sum_{n=1}^{\infty}(-)^nA_n\mu_n\right]. \qquad (77)$$

From (31), (32), (43) to (46), (65) to (68), (77), we deduce after considerable algebraic reduction that

$$U_r=\frac{2mr}{\Lambda\pi(1+\sigma)}\left[\frac{2m\pi(1+\sigma)}{(r^2+m^2)}\cos r\pi\sinh m\pi+\frac{2m^3\pi(1+\sigma)^2}{(r^2+m^2)^2}\cos r\pi\sinh m\pi\right.$$
$$+\frac{1}{(r^2+m^2)}\{(3-\sigma)(1-\sigma)\sinh^2 m\pi-m^2\pi^2(1+\sigma)^2\}$$
$$\left.-\frac{2m^2(1+\sigma)(3-\sigma)}{(r^2+m^2)^2}\sinh^2 m\pi\right], \qquad (78)$$

$$V_r=\frac{2mr}{\Lambda\pi(1+\sigma)}\left[-\frac{2m\pi(1+\sigma)}{(r^2+m^2)}\sinh m\pi-\frac{2m^3\pi(1+\sigma)^2}{(r^2+m^2)^2}\sinh m\pi\right.$$
$$-\frac{\cos r\pi}{(r^2+m^2)}\{(3-\sigma)(1-\sigma)\sinh^2 m\pi-m^2\pi^2(1+\sigma)^2\}$$
$$\left.+\frac{2m^2(1+\sigma)(3-\sigma)}{(r^2+m^2)^2}\cos r\pi\sinh^2 m\pi\right], \qquad (79)$$

where
$$\Lambda=(3-\sigma)^2\sinh^2 m\pi-m^2\pi^2(1+\sigma)^2.$$

From (78) and (79) it is at once apparent that

$$U_r=(-)^{r+1}V_r. \qquad (80)$$

We now express the left-hand side of equation (69) in terms of sines of multiples of y, whence equating to zero the coefficient of $\sin ry$ we deduce

$$(r^2+m^2)^2A_r+m^4\sum_{n=1}^{\infty}e_{rn}A_n-4m^2\sum_{n=1}^{\infty}f_{rn}A_n-imQ\sum_{n=1}^{\infty}c_{rn}A_n$$
$$-imQ\sum_{n=1}^{\infty}g_{rn}A_n+\frac{PRm^4A_n}{(r^2+m^2)^2}+PR\left[L_r\left\{\sum_{n=1}^{\infty}A_n(\alpha_n+\gamma_n)\right\}\right]$$
$$+M_r\left\{\sum_{n=1}^{\infty}(-)^nA_n(\alpha_n+\gamma_n)\right\}+U_r\sum_{n=1}^{\infty}A_n\mu_n-V_r\sum_{n=1}^{\infty}(-)^nA_n\mu_n$$
$$+PR\left[\sum_{n=1}^{\infty}e_{rn}A_n+\frac{4}{m^2}\sum_{n=1}^{\infty}f_{rn}A_n\right]=0,$$

an infinite system of equations analogous to (52).

In order that a non-zero solution of these equations should exist, it is necessary that the infinite determinant formed by eliminating the A's should vanish. This gives us the following determinantal equation from which to determine the critical value of S:

$$\begin{vmatrix} \Xi_1 & -imQ(c_{12}+g_{12}) & N_{13} & -imQ(c_{14}+g_{14}) \\ -imQ(c_{21}+g_{21}) & \Xi_2 & -imQ(c_{23}+g_{23}) & N_{24} \\ N_{31} & -imQ(c_{32}+g_{32}) & \Xi_3 & -imQ(c_{34}+g_{34}) \\ -imQ(c_{41}+g_{41}) & N_{42} & -imQ(c_{43}+g_{43}) & \Xi_4 \\ \cdots & \cdots & \cdots & \cdots \end{vmatrix} = 0, \qquad (81)$$

where
$$\Xi_r = (r^2+m^2)^2+m^4e_{rr}-4m^2f_{rr}+\frac{PRm^4}{(r^2+m^2)^2}+PR\left\{e_{rr}+\frac{4f_{rr}}{m^2}\right\}$$
$$+PR\{(\alpha_r+\gamma_r)(L_r+(-)^rM_r)+\mu_r(U_r-(-)^rV_r)\}, \qquad (82)$$

$$N_{ij} = m^4e_{ij}-4m^2f_{ij}+PR\left\{e_{ij}+\frac{4f_{ij}}{m^2}\right\}$$
$$+PR\{(\alpha_j+\gamma_j)(L_i+(-)^jM_i)+\mu_j(U_i-(-)^jV_i)\}. \qquad (83)$$

It is now necessary to consider the convergence of the infinite determinant involved in the equation for S, and for this purpose it is convenient to summarize the orders of magnitude of the various quantities involved. We have

$$\left.\begin{array}{ll} \alpha_r = O\left(\frac{1}{r}\right), & \gamma_r = O(r), \\ L_r = O\left(\frac{1}{r}\right), & \mu_r = O(r), \\ M_r = O\left(\frac{1}{r}\right) & e_{rn} = O\left(\frac{n}{r^3}\right), \\ U_r = O\left(\frac{1}{r}\right), & f_{rn} = O\left(\frac{n}{r}\right), \\ V_r = O\left(\frac{1}{r}\right), & g_{rn} = O\left(\frac{n}{r}\right). \end{array}\right\} \qquad (84)$$

Divide the rth row by r, and the rth column by r^3, then from (82) to (84) it follows at once that the product of the diagonal terms converges absolutely, and that the sum of the non-diagonal terms not involving the c's also converges absolutely. The terms involving the c's we consider separately. The sum of their modulus is

$$Qm \sum_{r=1}^{\infty} \sum_{n=1}^{\infty} \frac{|c_{rn}|}{rn^3},$$

which is less than

$$\frac{2Qm}{\pi}\left\{\sum_{r=1}^{\infty}\frac{1}{r^2}\right\}^2 + \frac{2Qm}{\pi}\sum_{r=1}^{\infty}\sum_{n=1}^{\infty}\frac{1}{rn^2\,|r-n|}, \quad [r \neq n]$$

which is convergent by virtue of (57) on interchanging the suffixes r and n. Since the determinant is convergent, we now approximate as in Case I.

7—In order to reduce the algebra and calculation involved to a minimum, we go straight to a *second* approximation formed by taking the first nine elements of the determinant in question (81). After certain simplification this reduces to

$$S = \frac{D\pi^2}{2mb^2}\left[\frac{\Xi_1\Xi_2\Xi_3 - \Xi_2 N_{13} N_{31}}{\begin{array}{r}\{N_{13}(c_{21}+g_{21})(c_{32}+g_{32}) + N_{31}(c_{12}+g_{12})(c_{23}+g_{23}) \\ -\Xi_1(c_{23}+g_{23})(c_{32}+g_{32}) - \Xi_3(c_{12}+g_{12})(c_{21}+g_{21})\}\end{array}}\right]^{\frac{1}{2}}.$$

The main results of the subsequent calculation are given in Table II. They are shown graphically in figs. 2 and 3.

TABLE II

	Second approximation	
$\sqrt{RP}$	m^2	k
0	1·48	1·664
1	1·55	1·717
2	1·75	1·855
3	1·95	2·035
4	2·1	2·245
5	2·2	2·471

Here k has the same meaning as in Case I, and it is interesting to note that for a plane strip the value of k obtained by Southwell and Skan is 1·68.

8—Before drawing any deductions from the results we have obtained, it is first necessary to consider to what extent our third approximation in Case I, and our second approximation in Case II, are likely to differ from the true results.

In Case I we compare the results obtained in this paper for the special case when the strip is flat with those obtained by Southwell and Skan (1924). In the first approximation there is an error of some 5 %, while in the second and third approximations this is reduced to a fraction of 1 %. For the general case in which the plate is curved, no comparison with any known result is possible, but by a consideration of the relative accuracy of the three approximations when the strip is flat, it seems probable that for the larger values of RP considered, the error in the third approximation will not exceed 1 %, while for the smaller values it will certainly be less.

In Case II it is not so easy to ascertain the error in the second approximation, but by a comparison with the results of Southwell and Skan for this case, it is clear that the second approximation is not as accurate as in Case I. In spite of this however the results of Case II, as shown in fig. 1, are quite consistent with those of Case I, the curves representing the third approximation in Case I and the second approximation in Case II tending to become parallel.

With reference to the possibility of obtaining further approximations in the two cases, there is no theoretical difficulty, but in view of the calculation which would be necessary and in view of the fact that the results wanted have already been obtained with an accuracy sufficient for all practical purposes, the labour involved would not appear to be worth while.

On considering the detailed results of the calculation (not actually given in this paper), it can be seen that for a given value of RP, a small change in m^2 from its critical value will give rise to a *very* much smaller change in k. Hence in view of the accuracy to which the results have been evaluated, m^2 is only accurate to two figures, and the third figure, where given, is liable to an error of anything up to three.

A final point to which it would appear appropriate to draw attention is the importance and use of Fourier series in solving equations of the type with which we have had to deal in this paper.

In conclusion, the writer wishes to thank Mr. W. R. Dean, Trinity College, Cambridge, for his valuable criticism and advice.

SUMMARY

A long plate, slightly curved, is subjected to a shearing force applied uniformly along its two straight edges. A theoretical investigation is carried out to ascertain the value of the shearing force at which buckling of the plate may be expected to take place.

Two cases are investigated, in one of which the two straight edges are simply supported, and in the other of which they are clamped. The results of the investigation, giving the critical shearing force in terms of the dimensions of the plate in each of the two cases considered, is shown graphically in fig. 2.

REFERENCES

Carslaw, H. S. 1930 "Introduction to the Theory of Fourier's Series and Integrals", 3rd ed.
Cox, H. L. 1933 *Rep. Memor. Aero. Res. Comm.*, No. 1553.
Dean, W. R. 1925 *Proc. Roy. Soc.* A, **107**, 734.
— 1926 *Proc. Roy. Soc.* A, **111**, 144.
Love, A. E. H. 1927 "Mathematical Theory of Elasticity", 4th ed.
Southwell, R. V. and Skan, S. W. 1924 *Proc. Roy. Soc.* A, **105**, 582.
Timoshenko, S. 1934 "Theory of Elasticity", 1st ed. McGraw-Hill.
— 1936 "Theory of Elastic Stability", 1st ed. McGraw-Hill.
Whittaker, E. T. and Watson, G. N. 1927 "Modern Analysis", 4th ed.

The Absorption Spectra of the Monosulphides of Alkaline Earth Elements and their Latent Heats of Vaporization

By L. S. Mathur, M.Sc.

Department of Physics, University of Allahabad, Allahabad, India

(*Communicated by M. N. Saha, F.R.S.—Received 6 April 1937*)

[Plates 2, 3]

INTRODUCTION

An important extension of Franck's ideas regarding the photo-dissociation of alkali-halides was made by P. K. Sen-Gupta and others of this laboratory. Sen-Gupta postulated that in compounds like ZnO, CdO, HgO, ZnS, CdS and HgS the chemical binding is ionic in nature so that the molecule has the structure: $M^{++}X^{--}$, where M and X are respectively the electro-positive and electro-negative elements. In a number of papers Sen-Gupta emphasizes the view that the primary processes of absorption of light by this class of compounds results in the simultaneous transition of two electrons from the electro-negative to the electro-positive atom, so that the molecule dissociates into two neutral atoms in their ground states. The photo-chemical equation can be represented as given below:

$$MX + h\nu_1 = M + X(^3\mathrm{P}), \tag{1}$$

where ν_1 is the frequency of light capable of bringing about the dissociation of the molecule. This will give rise to an absorption cut towards the long wave-length side of the continuous spectrum.

It may happen that besides dissociating into neutral normal atoms, one of the dissociated atoms may come out in an excited state as a result of the absorption of light of higher frequency. The probability of getting an excited atom is enhanced a good deal in those cases where the electronegative element possesses a few metastable levels. This is exactly the case for oxides and sulphides. It is well known that oxygen and sulphur atoms possess the metastable spectroscopic levels $^1D_2$ and $^1S_0$ besides the normal ^3P levels, so that the other two secondary processes of photo-dissociation can be represented by equations similar to (1). We may have, in fact,

$$MX + h\nu_2 = M + X(^1D), \qquad (2)$$

$$MX + h\nu_3 = M + X(^1S). \qquad (3)$$

If the processes (2) and (3) actually take place, we should expect the appearance of a second and third cut in the absorption spectrum followed of course by intermediate retransmissions of light. The correctness of the above postulate is well borne out by the following two important facts:

(1) The beginning of the first absorption is found to correspond with the calculated value as obtained from (*a*) thermochemical data; (*b*) determination of energy of formation of the crystal lattice out of ions.

In both of these methods we calculate the atomic heat of dissociation R, i.e. the energy required to break the molecule into its normal atomic constituents, where

$$R = \frac{N.h\nu}{J}. \qquad (4)$$

Knowing R, the beginning of absorption in frequency units can be determined at once with the help of equation (4).

(*a*) *Thermochemical Method*—The following thermochemical equations

$$\begin{aligned}
[M] &= M - L_M, \\
[S_2] &= S_2 - 2L_S, \\
S_2 &= 2S - D_{S_2}, \\
[MS] &= MS - L_{MS}, \\
[M] + \tfrac{1}{2}[S_2] &= [MS] + Q, \\
M + S &= MS + R,
\end{aligned}$$

can be linked with each other to yield equation (5),

$$R + L_{MS} = Q + L_M + L_S + \tfrac{1}{2}D_S\,. \qquad (5)$$

The process can be represented by the $\triangle ABD$ in fig. 1, when we consider the transition of the solid constituents of the sulphide (A) to their gaseous state (D) via the solid sulphide (B). The cycle is the well-known Born's cycle. The three sides of the $\triangle ABD$ yield the equation (5), where

Q = heat of formation of the solid sulphide out of solid metal $[M]$ and sulphur $[S]$,
D_{S_2} = heat of dissociation of sulphur S_2 into S atoms,
L_M = heat of vaporization of the metal $[M]$,
L_S = heat of vaporization of sulphur $[S]$,
L_{MS} = heat of vaporization of the compound $[MS]$.

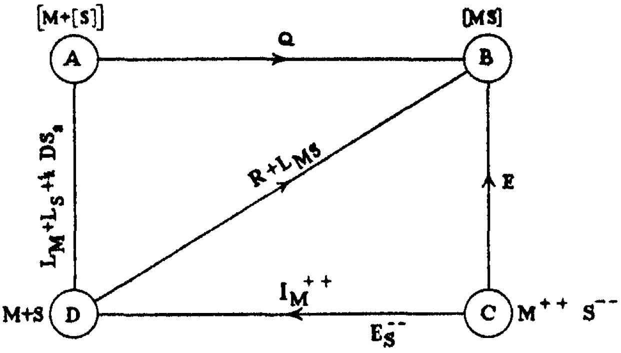


FIG. 1—Born's cycle.

In the present investigation, however, the L_{MS} values for the monosulphides of alkaline earths are not known, and therefore the position of the first cut towards the long wave-length side cannot be calculated beforehand from equation (5), so as to check its validity from subsequent data.

It is customary to get an approximate value of the latent heat of the compound with the aid of Trouton's rule, in cases where its value is unknown, due to lack of vapour-pressure data. We have

$$\frac{L_{MX}}{T} = \text{constant}, \tag{6}$$

where T is the boiling-point of the compound in absolute temperature scale. It is rather unfortunate that even the boiling-points of these refractory compounds which I have undertaken to examine spectroscopically are not known. I have, therefore, calculated the latent heats of vaporization of the monosulphides, taking equation (5) to be true as shown by various workers (Dutta 1932; Sen-Gupta 1933, 1934*a*, 1934*b*; and Sharma 1933).

(*b*) *Grating Energy Method*—Let E denote the energy required to form the crystal MX out of the ions M^{++} and X^{--}, so that

$$M^{++}+2e = M+I_{M^{++}}. \tag{7}$$

Similarly, we have $$S^{--}-2e = S-E_{S^{--}}, \tag{8}$$

where $I_{M^{++}}$ is the sum of ionization potentials of M and M^{+}, and $E_{S^{--}}$ is the heat of formation of an S^{--} ion out of S and two electrons. From the $\triangle BCD$ of Born's cycle (fig. 1) we get

$$E = R+L_{MS}+I_{M^{++}}-E_{S^{--}}, \tag{9}$$

$$R = E-L_{MS}-I_{M^{++}}+E_{S^{--}}. \tag{10}$$

The lattice energy E in equation (10) can be found by Born and Gerlach's formula

$$E = K.\frac{n-1}{n}\sqrt[3]{(\rho/M)}, \tag{11}$$

where K = constant involving Madelung's coefficient,
ρ = density,
M = molecular weight,
n = repulsive exponent and is given by

$$n = 1+C.\frac{1}{\chi}\left(\frac{M}{\rho}\right)^{\frac{4}{3}}, \tag{12}$$

where C = constant depending upon crystal structure,
χ = compressibility.

The formula given above is derived with the help of electrostatic grating theory. Recently a new formula for the grating energy has been given in which besides the Coulombian part of energy, the contributions of energy due to Van der Waals' cohesive force, null-point oscillation energy and the repulsive potential (with two arbitrary constants and an exponential term) have also been taken into account (Born and Mayer 1932). The complete formula with different parts of energy in the order given above is of the form

$$\Phi(r_0) = -\frac{\alpha e^2}{r_0}+\frac{C}{r_0^6}+\epsilon-(1-k)\frac{\rho}{r_0}\left\{\frac{\alpha e^2}{r_0}+\frac{6e}{r_0^6}-\frac{3VT}{N}\left(\frac{\partial p}{\partial T}\right)_V\right\}, \tag{13}$$

where $\Phi(r_0)$ = grating energy,
α = Madelung number,
e = elementary charge,
r_0 = grating interval at T = 273° A,

$\epsilon = \frac{9}{4} h . \nu_{\text{max.}}$,

$\nu_{\text{max.}}$ = maximum frequency of vibration,

$$k = \frac{(a-1)\, M' C_2 \left(1 + \frac{C_1}{C_2} e^{-2\delta/\rho}\right) . e^{-(\delta/r_0+1-a) r_0/\rho}}{2M + aM' C_2 \left(1 + \frac{C_1}{C_2} e^{-2\delta/\rho}\right) . e^{(\delta/r_0+1-a) r_0/\rho}},$$

M = number of different kinds of ions in the grating,

M' = number of the same kind of ions in the grating,

$C_1 = C_{++}$,

$C_2 = C_{--}$,

$C = S_6' C_{+-} + S_6'' \frac{C_{++} + C_{--}}{2}$,

δ = difference of ionic radii,

ρ = universal constant,

$S_6' = \Sigma \frac{1}{R^6}$ for different kinds of grating points,

$S_6'' = \Sigma \frac{1}{R^6}$ for the same kind of grating points,

C_{+-} = interaction constant between positive and negative ion,

C_{++} = interaction constant between positive and positive ion,

C_{--} = interaction constant between negative and negative ion.

Soon after the announcement of the new grating theory, the grating energy of alkali halides and the electron affinity of halogens were calculated with the help of the new formula and found to be in agreement with observed data (Mayer and Helmholtz 1932). The grating energy of the oxides and sulphides of alkaline earths were calculated by Born's formula (Mayer and Maltbie 1932). In the present investigation their values for the monosulphides of Ca, Sr and Ba are alone of interest. In Table I are given the results of their calculations.

TABLE I

Salt	Coulombian energy Ergs/ mol. × 10^{12}	Van der Waals' energy Ergs/ mol. × 10^{12}	Null-point oscillation energy Ergs/ mol. × 10^{12}	Repulsive potential energy Ergs/ mol. × 10^{12}	Total energy Ergs/ mol. × 10^{12}	Total grating energy
CaS	−57·5	−0·2	0·1	6·8	−50·9	−737
SrS	−53·3	−0·2	0·1	6·0	−47·4	−686
BaS	−49·8	−0·2	0·1	5·3	−44·6	−647

The other factors in equation (10) can be considered now. The value of $I_{M^{++}}$ as obtained from spectroscopic data are 441·7, 383·0 and 348 kcal. for Ca, Sr and Ba respectively. The quantity $E_{S^{--}}$, the electron affinity of sulphur, has been given by Born and Gerlach, Samuel and Lorentz, and Mayer and Maltbie, but there is no agreement in these values, and they vary from 30 to -80 kcal. Mayer's value has been obtained from the $\square\,ABCD$ of Born's cycle as shown in fig. 1, with the help of grating energy given in Table I. From equation (10) we could calculate R provided L_{MS} and $E_{S^{--}}$ are known. Since L_{MS} is not known, instead of calculating R we could calculate L_{MS} from the observed data regarding R. Even this cannot be done so long as a reliable value of the electron affinity of sulphur is not known. We can now pass on to the discussion of the other points in support of the postulates made in the beginning of the article.

(2) Sen-Gupta observed a first and second cut in the case of ZnS, CdS and a third cut as well in the case of HgS, while for others he showed that the positions of the second and third cuts are beyond the region of observation. The differences $h(\nu_1-\nu_2)$ and $h(\nu_2-\nu_3)$ correspond to the $^3\mathrm{P}-{}^1\mathrm{D}$ and $^1\mathrm{D}-{}^1\mathrm{S}$ differences of the X atom respectively.

APPARATUS

On account of the highly refractory nature of the sulphides of alkaline earth elements, the vacuum graphite furnace of this laboratory which is capable of giving a temperature of about 2000° C. was employed. From the year 1927 when Saha Sur and Majumdar first carried out an experimental verification of Saha's ionization theory, the vacuum graphite furnace designed on the lines indicated by Professor Arthur King of Mt. Wilson Observatory has undergone a series of modifications (Desai 1932), and experience has led us to a further modification of our furnace which I shall now describe. It will, however, be noted that the salient features of the furnace are not very different from the original.

In fig. 2 is given a diagrammatic sketch of the furnace together with a few of its accessories. A water-cooled cast-iron base carries two copper electrodes EE which are also water-cooled and are rigidly clamped to the base by means of bolts as shown in fig. 2. At the ends of the electrodes are brazed two copper plates 10 × 7 cm. Two graphite plugs FF which carry the current-bearing graphite tube G can be bolted at different positions on the copper plates P. This method of mounting the graphite tube permits the variation in the effective length of the heating tube. The maximum length of the tube G between the plugs F is about 30 cm. in the present investigation in contrast with much smaller length used in the previous investigations. The long tube gives a long column of the vapour of the substance which is being vaporized and consequently affords much better conditions of absorption. On the other hand, it possesses the disadvantage that the maximum temperature attained with the limited amount of power supply at our disposal is not more than 1800° C. Over the base *B* is placed the main body of the furnace which is also double-walled and rectangular in shape. This double-walled body is provided with four windows *W* on the four sides.

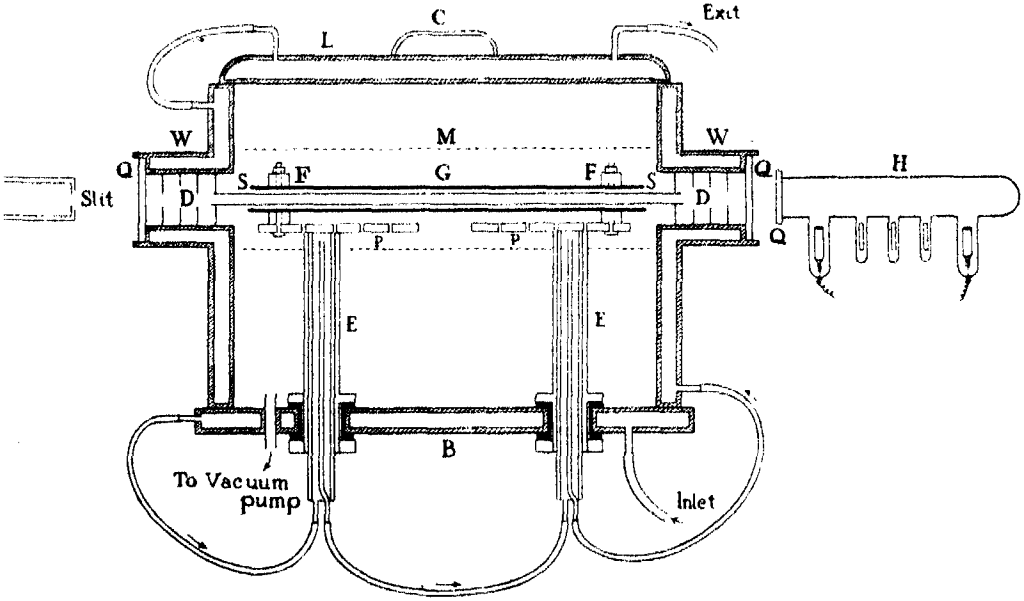


FIG. 2—Vacuum furnace.

The windows facing the hydrogen discharge tube and the quartz spectrograph are closed by two quartz plates *QQ*. To the other two windows only glass plates are attached. One of these two windows is used for visual observation of the graphite tube or to record its temperature by a disappearing filament pyrometer. Through the other we can either make electrical connexions for thermocouples, etc., or a discharge tube can be connected to indicate vacuum. Behind the quartz plates *QQ* are placed brass diaphragms to prevent the condensation of the vapour over them. The furnace is closed by a water-cooled lid *L*, provided with a handle *C* which facilitates easy cleaning of the furnace. The water flows first into the lower plate, and after washing the electrodes enters the main body from where it is led into the lid and finally goes to the sink from the top. To diminish the heating of the body as a result of the intense radiation from the heated graphite tube, an iron mantle *M* is placed surrounding the graphite tube with a hole just opposite the observational window. This arrangement gives a little higher temperature, for the radiation loss is prevented a good deal.

EXPERIMENT

While taking the absorption spectra of the monosulphides of Ca, Sr and Ba, the substances were placed in a quartz tube and a heavy current was passed through the surrounding graphite tube from the terminals of a low-tension welding transformer, which gives a maximum of about 1000 amp. at a pressure of about 10 V. The furnace was evacuated by a vacuum pump, and then nitrogen from a cylinder was introduced at known pressures. For ultra-violet region E_2 and E_{31} quartz spectrographs were used, while for the visible a constant deviation glass spectrograph was employed. Hydrogen discharge tube run by a 2 kW transformer gave sufficiently intense light in the ultra-violet. A Philips coiled-coil 40 W lamp served as a source of continuous light for the visible region. Different experimental conditions under which each salt gave a suitable absorption spectrum are given below for individual cases.

RESULTS

CaS—Even up to a temperature of 1000° C. the salt did not show any continuous absorption. Only sulphur bands of Christy and Naudé appeared strongly on the plates, showing that the sulphide was contaminated with excess of free sulphur. Merck's extra-pure compounds were found to smell of sulphur and H_2S, when freshly introduced in silica tubes. At times when the current heating the furnace was suddenly withdrawn the spectrograms showed the presence of far ultra-violet bands of sulphur lying beyond λ 2200, which are obtained at low temperatures. These bands as usual were found to be rather diffuse. When a temperature of 1150° C. was exceeded, the sulphide of calcium melted and gave a genuine absorption beginning at λ 2388. The long wave-length limit of the continuous absorption was located by taking a microphotogram of the plate. The spot of light was run on the same plate twice, first on the continuous spectrum, and then on the absorption spectrum. The position of the cut was measured by using Hartmann's dispersion formula. The peaks shown in fig. 5, Plate 2, corresponding to a few lines in the spectrum of copper, are used for comparison.

SrS—Without filling the furnace with nitrogen, no absorption due to the vapour of strontium sulphide could be obtained right up to a temperature of 1200° C. on account of the diffusion of the salt vapour in the vacuum. Beyond 1200° C., however, the salt gave enough vapour to show continuous absorption even in vacuum. Nitrogen was filled in the furnace and the absorption spectrum was photographed at the lowest temperature possible. This temperature was about 1150° C. Below this temperature only the sulphur bands, mentioned above, were obtained on the plates. On the large E_2 quartz spectrograph only one cut appeared on the plate. By decreasing the intensity of continuous light and employing a smaller E_{31} spectrograph of higher light-gathering power two cuts could be easily photographed. The microphotogram of the absorption spectrum given in fig. 6, Plate 2, shows the positions of these cuts distinctly. The two cuts were located at λ4500 and λ3566.

BaS—Like the sulphides of calcium and strontium, this also gave at low temperatures only sulphur bands. Two regions of absorption were observed in the absorption spectrum of this sulphide which on measurements were found to be situated at λ4441 and λ3515. The absorption in the visible region was photographed on a glass spectrograph, while the cut in the ultra-violet was taken on a quartz spectrograph. The two microphotometric curves for the salt are shown in figs. 7 and 8, Plate 3. In Table II are given the positions of different beginnings of absorptions from long wave-length limits in A as well as kcal.

TABLE II

	1st absorption		2nd absorption	
Salt	A	kcal.	A	kcal.
CaS	2388	119·76	—	—
SrS	4500	63·55	3566	80·20
BaS	4441	64·39	3515	81·36

CALCULATIONS

As pointed out earlier in the introduction the equation (5) is employed to determine the latent heats of vaporization of the salts. The thermochemical data necessary for the calculation of L_{MS} is collected in Table III from Landolt-Bornstein tables.

TABLE III

Salt	Q kcal.	$\frac{1}{2}D_{S_2}$ kcal.	L_M kcal.	L_S kcal.	R kcal.	L_{MS} kcal.
CaS	92·0	51·3	44·2	14·65	119·76	82·39
SrS	99·2	51·3	32·5	14·65	63·55	134·10
BaS	102·5	51·3	32·4	14·65	64·39	136·46

In Table III the value of heat of dissociation of S_2 into S atoms is that from the determination of Christy and Naudé (1931) from the dissociation of sulphur bands. The last column gives the value of L_{MS} of the sulphides obtained from the observed value of R and the thermochemical data of the salts.

INTERPRETATION OF RESULTS

In the present investigation two regions of continuous absorption with an intermediate patch of transmitted light have been obtained for SrS and BaS, while in the case of CaS only one region of absorption was observed. The second absorption for CaS would lie beyond the region of a quartz spectrograph. From the positions of the long wave-length limits of continuous absorption it is easily seen that $h\nu_2 - h\nu_1$ for the sulphides of Sr and Ba are as given below:

$$\mathrm{SrS}(h\nu_2 - h\nu_1) = 16{\cdot}65\,\text{kcal.} = 0{\cdot}722\,\text{e-volt.}$$

$$\mathrm{BaS}(h\nu_2 - h\nu_1) = 16{\cdot}97\,\text{kcal.} = 0{\cdot}736\,\text{e-volt.}$$

Due to the absorption of light of frequency ν_1 the sulphide MS is split up into two normal atoms M and S (^{3}P). But when light of higher frequency ν_2 is absorbed by the molecule the sulphur atom may be excited to the next higher state $^1D_2$. If this scheme of dissociation is true, then the difference in the long wave-length limits of the two continuous absorptions should correspond to the atomic term difference $^3P - {}^1D_2$ of sulphur. Absolute values of triplet terms in the arc spectrum of sulphur were first obtained by Frerichs (1933), but the singlet terms were added later (Ruedy 1933). The difference $^3P - {}^1D_2$ is as follows:

$$\begin{aligned} {}^3P - {}^1D_2 &= 83554 - 74315 \\ &= 9239\,\text{cm.}^{-1} = 1{\cdot}14\,\text{e-volts.} \end{aligned}$$

The difference $h\nu_2 - h\nu_1$ for the two regions of continuous absorption obtained here is apparently smaller than the atomic term difference obtained from the arc spectrum of sulphur. A discrepancy of this nature has been observed by all earlier workers for different compounds. Thus, in the case of SO_2 it was found that $^3P - {}^1D_2$ difference was 1·65 e-volts, while the analysis of arc spectrum gives a value of 1·95 e-volts (Dutta 1932). From the absorption spectrum of N_2O (Sen-Gupta 1934*d*) also, a lower value of the term difference $^4S - {}^2D$ of nitrogen was obtained; while, in another case, the same author (Sen-Gupta 1932) obtained a higher value of 1·3 e-volts for the interval $^3P - {}^1D_2$ of sulphur from absorption experiments. A suitable interpretation of these discrepancies has already been given (Sen-Gupta 1934*c*). The same arguments may be applied to the absorption spectra of the monosulphides of alkaline earth elements by considering the Frank-Condon energy diagrams.

We assume that in the normal state of the molecule the potential energy curve has a deep minimum, while the curve for the upper unstable state has no minimum. By absorption of light, the molecule is raised from the normal to the unstable repulsive state. The type and characteristics of absorption depend upon the form of the $(U.r)$ curve for the unstable state. A steep curve would give rise to gradual absorption, while a flat curve would give sharp absorption.

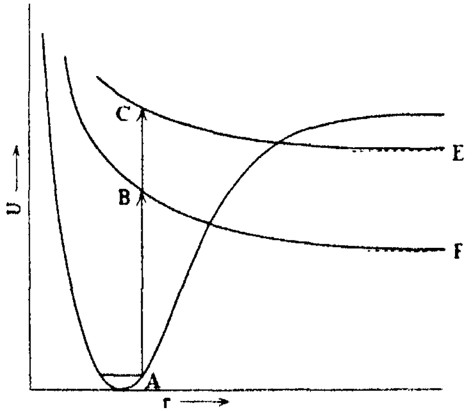


FIG. 3

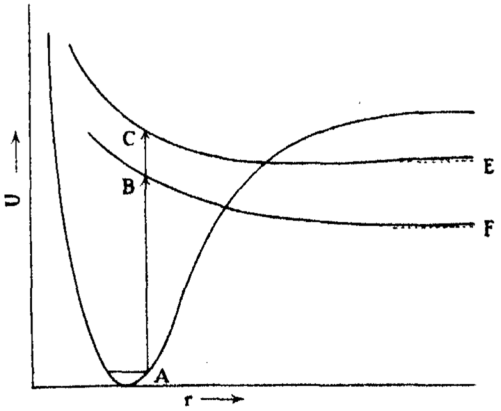


FIG. 4

In accordance with the theory we find that for BaS the first cut is very gradual as shown in fig. 7, Plate 3, but the second cut is fairly sharp. The $(U.r)$ curves for these two unstable states in the case of BaS would therefore be as shown in fig. 3. The frequency difference $h\nu_2 - h\nu_1$ corresponds to BC as obtained from the absorption experiments, while actually it ought to be equal to EF. But $BC < EF$ and consequently a lower value is obtained than expected from atomic term difference.

In the case of SrS, on the other hand, it appears from the microphotometric curves that the second cut is not so sharp as the first one, but since the dispersions in regions of two absorptions are different the microphotograms are not expected to give the true nature of absorption. On plotting the percentage of absorption against wave-length it was found that the two absorptions are equally sharp, and therefore the upper state curve takes the shape shown in fig. 4.

Here again, the difference *BC* is less than *EF*, and therefore the difference $h\nu_2 - h\nu_1$ is lower than the expected value. Thus the slope of the upper state curve as determined from the microphotograms and the percentage of absorption-wave-length curves helps us to explain the apparent discrepancy of the observed results.

Finally, I feel it a great pleasure to express my most cordial thanks to Professor M. N. Saha, F.R.S., for his valuable help and guidance in the present work. I am indebted to Professor A. T. Mukerjee and Professor Kamta Prasand of Patna Science College for allowing me the use of their microphotometer, and to Mr. B. N. Ghosh for the help which he gave me in working with the instrument. I am also thankful to Dr. P. K. Sen-Gupta for a discussion of some of the points in the paper.

SUMMARY

The paper contains a study of absorption spectra of the sulphides of alkaline earth elements.

Two regions of absorption have been reported for SrS and BaS, while for CaS only one region of absorption in far ultra-violet has been observed.

From the long wave-length limits of the first continuous absorption latent heats for CaS, SrS and BaS have been found to be 82·39 kcal., 184·10 kcal. and 136·46 kcal. respectively.

The value of $^3P - {}^1D_2$ term difference for sulphur has been found to be 0·72 e-volt, which is less than 1·14 e-volts obtained from atomic spectrum. It has been explained with the help of Frank-Condon energy diagrams.

REFERENCES

Born, M. and Mayer, J. E. 1932 *Z. Phys.* **75**, 1–18.
Christy, A. and Naudé, S. M. 1931 *Phys. Rev.* **37**, 903–19.
Desai, M. S. 1932 *Proc. Roy. Soc.* A, **136**, 76–82.
Dutta, A. K. 1932 *Proc. Roy. Soc.* A, **137**, 366–72.
Frerichs, R. 1933 *Z. Phys.* **80**, 150–66.
Mayer, J. E. and Helmholtz, L. 1932 *Z. Phys.* **75**, 19–29.
Mayer, J. E. and McC. Maltbie, M. 1932 *Z. Phys.* **75**, 748–52.
Ruedy, J. E. 1933 *Phys. Rev.* **43**, 1045.
Sen-Gupta, P. K. 1932 *Proc. Roy. Soc.* A, **138**, 84–91.
— 1933 *Bull. Acad. Sci. Allahabad*, **3**, 65–8.
— 1934*a* *Proc. Roy. Soc.* A, **143**, 438–54.
— 1934*b* *Bull. Acad. Sci. Allahabad*, **3**, 203–8.
— 1934*c* *Z. Phys.* **88**, 647–60.
— 1934*d* *Bull. Acad. Sci. Allahabad*, **3**, 197–202.
Sharma, R. S. 1933 *Bull. Acad. Sci. Allahabad*, **3**, 17–22.

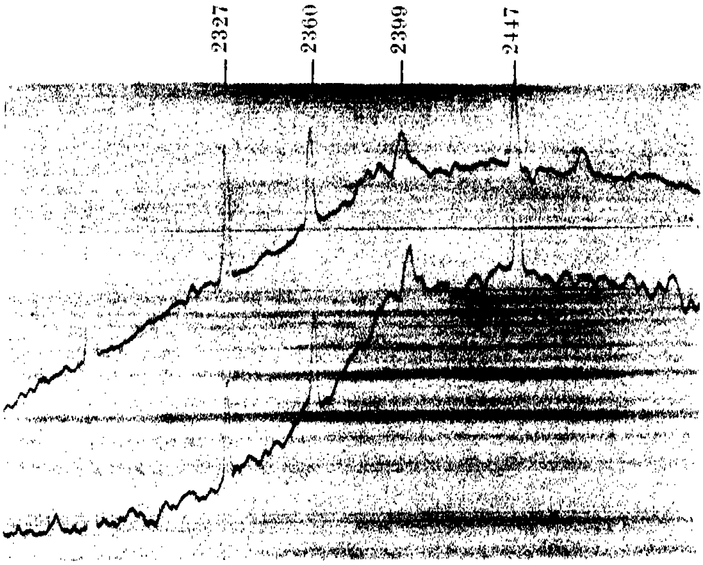


FIG. 5—Microphotogram for CaS absorption spectrum.

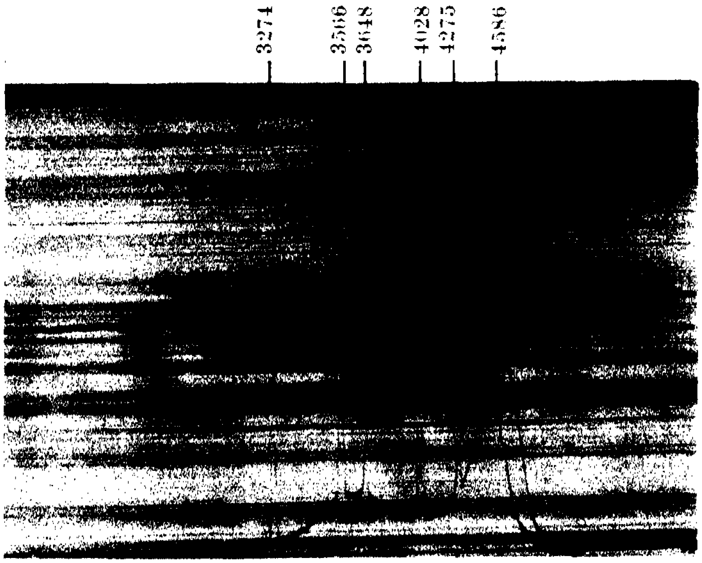


FIG. 6—Microphotogram for SrS absorption spectrum.

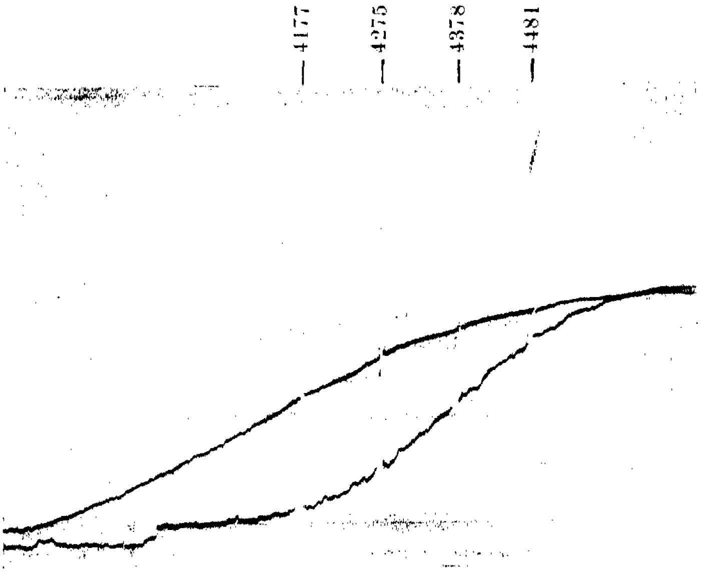


FIG. 7—Microphotogram for BaS absorption spectrum in the visible.

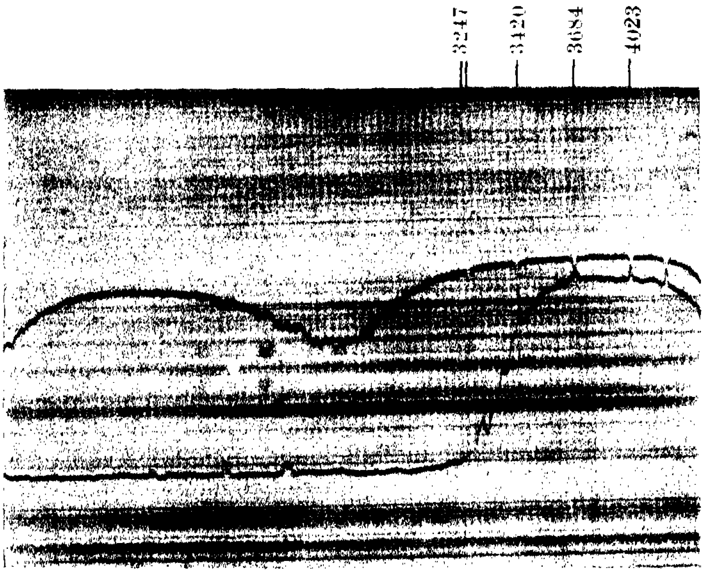


FIG. 8—Microphotogram for BaS absoption spectrum in the ultra-violet.

The Angular Distribution of Photoelectrons from the K Shell

By F. F. H. Eggleston, B.Sc., and L. H. Martin, Ph.D.
Natural Philosophy Department, University of Melbourne

(*Communicated by T. H. Laby, F.R.S.—Received* 21 *April* 1937)

Introduction

In an investigation with the cloud chamber on the properties of X-rays, Wilson (1923) observed that although photoelectrons due to a narrow pencil of X-rays were ejected from the gas atoms in all directions, there was, especially in the case of high-energy electrons, a preponderance with forward components in their velocities of ejection. Wilson's pictures threw light on results obtained earlier by workers in X-ray fields, who found an excess of electrons ejected forwards from thin metal foils by X-rays, and have since inspired a number of experiments designed to determine with precision the angular distribution of photoelectrons. A survey of the different methods developed for this purpose has been given by Compton and Allison (1935). Of these the cloud expansion chamber has proved the most satisfactory, in spite of the inevitable disadvantages associated with statistical methods. If a suitable gas is chosen, the shells from which the electrons are ejected can usually be identified, and the directions of ejection can be determined from measurements relating to the initial portions of the cloud tracks. No trouble is ever encountered in distinguishing photoelectron tracks from those due to recoil electrons.

The method has been criticized by Watson (1927), who suggested the possibility of nuclear scattering within the radius of the first drop. This objection has been answered by Kirchner (1927), who was able to show from an application of Rutherford's theory of nuclear scattering to the problem, that under normal conditions of observation the effect of nuclear scattering on the initial directions of ejection as measured from track photographs could be neglected. The most likely source of error appears to lie in the systematic exclusion of tracks in certain angular intervals determined by the experimental arrangements (e.g. positions of cameras, lighting, etc.), and much thought has been given by different workers (Kirchner 1927; Williams, Nuttall and Barlow 1928) as to the best means for ensuring representative samples of tracks for measurement. In this respect our experience leads us to believe that a requirement of fundamental importance is sufficient and uniform illumination throughout the whole chamber.

It has become customary to define the distribution of photoelectrons in the following manner. If the OZ axis of a right-handed system of axes $OXYZ$ is the direction of the X-ray beam, and OP the direction of ejection of the photoelectron, then the angle ZOP is the longitudinal angle θ, and the angle which the plane ZOP makes with the plane XOZ is the azimuthal angle ϕ. Suppose $J(\theta, \phi)\,\delta\omega$ represents the number of photoelectrons whose directions have angular co-ordinates limited by a small solid angle $\delta\omega$ enclosing the direction OP. If there is axial symmetry about the X-ray beam the function $J(\theta, \phi)$ may be written $J(\theta)$, and the number ejected whose initial directions have longitudinal angles between θ and $\theta + d\theta$ will be

$$n(\theta)\,d\theta = 2\pi J(\theta)\sin\theta\,d\theta.$$

$n(\theta)$ is called the longitudinal distribution function, and it is the form of this which is sought in cloud chamber experiments. Actually the direct comparison of the experimental with the theoretical form of $n(\theta)$ does not constitute a sensitive test of theory, and it is customary to supplement this by the introduction of certain quantities defined below whose values depend upon the form of the distribution function.

The asymmetry of a distribution is the ratio of the number of electrons ejected in a forward direction to the number ejected in a backward direction:

$$\text{Asymmetry } A = \int_0^{\frac{1}{2}\pi} n(\theta)\,d\theta \Big/ \int_{\frac{1}{2}\pi}^{\pi} n(\theta)\,d\theta.$$

The bipartition angle θ_b is the angle such that the number of photoelectrons ejected at an angle less than θ_b with the beam is equal to the number ejected at an angle greater than θ_b with the beam, i.e.

$$\int_0^{\theta_b} n(\theta)\,d\theta = \int_{\theta_b}^{\pi} n(\theta)\,d\theta.$$

In quantum mechanics the problem of the distribution of photoelectrons is part of the larger one of the absorption of X-rays with the emission of photoelectrons, the forward bias of the paths of these being associated with the radiation pressure (Sommerfeld 1930). The usual method of approach follows the lines of perturbation theory in which the wave equation contains a term representing the perturbing effect of the electromagnetic field of the incident light. The solution contains not only discrete characteristic functions typical of the normal atom, but also a set of continuous characteristic functions with which is identified photoelectron emission, the current density at any point in space being a measure of the probability of ejection in that direction. Fischer (1931) obtained for the distribution function of K photoelectrons

$$J(\theta, \phi) \sim \frac{\sin^2\theta \cos^2\phi}{[1+k-\beta\cos\theta]^4},$$

where $k = h\nu/2mc^2$. The corresponding longitudinal distribution will be given by

$$n(\theta) \sim \frac{\sin^3\theta}{[1+k-\beta\cos\theta]^4}.$$

Quantum mechanics predicts a distribution of K photoelectrons almost completely determined by the velocity of ejection. The term involving $\beta = v/c$ appears because the electron receives forward momentum due to the magnetic force of the incident radiation. The expression above was derived from non-relativistic theory, but Sauter (1931) found that for values of β up to 0·5 relativistic calculations lead essentially to the same results, for example, the value of $\cos\theta_b$ differs by less than 1 % from the non-relativistic value.

Table I shows the results of experiments in which reasonable care was taken to ensure homogeneity of the X-ray beam and in which accurate angle measurements could be made. The experiments were all performed with cloud chambers. The results are arranged in order of increasing

TABLE I—DISTRIBUTION OF K PHOTOELECTRONS

Author	λ A	Atom	E_{kin} e-kV	β	Tracks	θ_b Exp.	θ_b Theor.	A Exp.	A Theor.
Anderson (1930)	0·59	Br	7·5	0·17	233	79°·0	80°·4	1·7	1·7
Auger (1928*a*)	0·71	A	14·2	0·24	1200	77°·0	76°·6	1·9	2·0
Loughridge (1927)	0·71	A	14·2	0·24	231	72°·0*	76°·6	3·5*	2·0
Loughridge (1927)	0·71	Air	17·0	0·26	123	72°·0*	75°·4	2·5*	2·2
Anderson (1930)	0·71	Air	17·0	0·26	272	74°·9	75°·4	2·2	2·2
Williams and others (1928)	0·60†	Air	20·0	0·28	779	75°·7	74°·1	2·1	2·3
Anderson (1930)	0·56	Air	22·0	0·29	200	73°·0	73°·5	2·4	2·4
Auger (1929)	0·21	Xe	24·3	0·31	600	76°·0	73°·0	—	—
Auger (1928*b*)	0·21	A	55·6	0·47	1000	69°·0	63°·9	3·3	4·2
Lutze (1931)	0·135	A	88·2	0·59	586	58°·0	57°·3	5·7	6·4
Auger and Meyer (1931)	0·130	A	88·8	0·60	?	64°·0‡	56°·6	—	—

* Estimated from the published distributions.

† The mean effective wave-length of several in the range 0·545–0·709 A.

‡ Value calculated from the result $S = 1{\cdot}45$ given in Auger's paper. He defines $S = \cos\theta_b/\frac{1}{2}\beta$.

value of β calculated non-relativistically, the values of the bipartition angle θ_b and asymmetry A obtained experimentally being tabulated, together with the theoretical values.

Most of the experimental results are for small values of β and show reasonable agreement with values of θ_b and A calculated from the wave mechanics distribution, but a steady divergence appears between the results of Auger and theory as β increases. Lutze, on the other hand, finds good agreement with theory for $\beta = 0{\cdot}59$. Special consideration should be given to the results obtained by Auger because each one was drawn from measurements on a large number of tracks; moreover, his experiments represent the only series covering a wide range of β values. The majority of other results have been obtained from approximately 200 tracks, which is barely enough to permit a close comparison of experiment with theory. It should also be mentioned that most of the authors have failed to state explicitly that their measurements were confined solely to electrons from the K shell. Wave mechanics predicts (Schur 1930) a symmetrical component in the distribution of electrons from the L shells, so that if L tracks are included in the measurements the values of θ_b will be greater than the expected values, while the values of A will be too small. No doubt the number of tracks usually measured is too small to show an appreciable effect over and above statistical variations.

The present investigation was undertaken in the attempt to throw some light on the questions raised by the consideration of Table I. A description is given here of measurements made with a cloud expansion chamber on photoelectrons ejected from the K shell of argon with energies 14·2, 46·8 and 88·2 e-kV respectively. Considerable difficulty is met in attempting energies much greater than 90 e-kV, since the number of photoelectrons is very small compared with the number of recoil electrons. Any attempt to obtain a reasonable number of photoelectron tracks leads to the confusion of their origins with large numbers of recoil tracks. With the highest energy photoelectrons we averaged less than two measurable tracks per plate.

EXPERIMENTAL PROCEDURE

The expansion chamber used in these experiments has been described by the authors (Martin and Eggleston 1937) in a previous paper on the Auger effect. The photography has been considerably improved by replacing the mercury arcs, previously used for illumination, with a mirror carbon arc running at 25–35 amp. A horizontal beam from the arc is passed through the chamber and reflected back by means of a plane mirror placed on the far side; in this way almost equal illumination is obtained for each camera. The exposure was timed by a roller-blind shutter working at 1/30 sec. placed between the chamber and the arc.* This form of illumination has three distinct advantages over the earlier type; the intensity of the illumination is sufficiently great to permit the lens apertures to be stopped down, thus giving greater depth of focus; the whole of the chamber is illuminated so that all tracks are of the same photographic intensity and are photographed from beginning to end; and finally the thermal effects associated with the heating elements of the mercury arcs are entirely eliminated. It is a great advantage to be able to stop down the apertures of the camera lenses in the present experiments as it helps to prevent the systematic choosing of certain groups of tracks whose directions are more favourable to photography.

The two cameras were mounted with their axes at right angles and intersecting at the centre of the chamber, the plane containing the axes being perpendicular to the X-ray beam. The lens apertures were usually $f/6{\cdot}3$, the lenses Zeiss Tessars of maximum aperture $f/4{\cdot}5$, and focal length 13·5 cm.

In making the necessary measurements the two plates were placed side by side in a vertical enlarger and enlarged pictures, three times the natural size, were projected on to a horizontal table. Each plate was adjusted in turn until a line ruled on the projection table coincided with the direction of the X-ray beam, indicated by the occurrence of track origins along it, then the angle between the positive direction of the X-ray beam and the initial direction of the track was measured. A further measurement was required, the distance of the track origin from the centre of the chamber, and this was obtained by a scale which was constructed to give the true distance directly in cm.

The tracks on the photographic plates are conical projections of the actual tracks, with the optical centres of the lenses as poles of projection. The longitudinal angle θ is given by

$$\tan^2\theta = \frac{(\cot\theta_x + \sigma)^2 + (\cot\theta_y + \sigma)^2}{(\cot\theta_x \cot\theta_y - \sigma^2)^2},$$

where θ_x, θ_y are the angles measured and $\sigma = d/l$, d being the distance of the track from the centre of the chamber and l the distance of the optical centres of the two lenses from the centre of the chamber.

---

* We are indebted to Professor T. H. Laby for suggesting this system of illumination.

If the track origin is at the centre of the chamber $\sigma = 0$ and the formula reduces to the form (orthogonal projection)

$$\tan^2\theta = \tan^2\theta_x + \tan^2\theta_y,$$

which is the form assumed by several investigators who confined their measurements to tracks near the centre of the chamber. With an expansion chamber of normal dimensions the assumption of orthogonal projection may lead to serious error; for example, with $\sigma = 0{\cdot}15$ the error in θ may be as great as eight degrees.

When the angles θ_x and θ_y and the distance d were measured on the projected pictures, the angle θ was found from graphs showing the variation of θ with θ_x and θ_y for different values of d. The use of a graphical method has definite advantages. It eliminates all chance of arithmetical slips, and since it is a quick and easy method, errors in the computation are reduced to a minimum. The elimination of subjective errors is an important factor in an experiment where very large numbers of tracks have to be measured. The measurement of the angles θ_x and θ_y on an enlarged picture has a distinct advantage, since the tracks are easily identified and the structure at the origin readily traced; moreover, the tracks can be chosen and discussed simultaneously by two observers.

In order to test the accuracy of the method used in computing the angle θ, artificial tracks were constructed from fine glass rods and photographed in different positions in the chamber. In Table II the true values of θ are compared with the values obtained by application of the orthogonal and conical formulae.

TABLE II—RESULTS OF MEASUREMENTS MADE ON GLASS TRACKS

d cm.	True θ	Orthogonal	Conical
−4·5	100·0	97·3	99·7
−3·75	36·5	36·7	36·3
−2·75	122·5	121·7	121·8
4·5	80	83·2	80·5
3·75	143·5	143·5	143·6
2·75	57·5	57·7	57·6

The necessity for the use of the exact conical theory is demonstrated by this table.

Homogeneous beams of X-rays were obtained in the first two experiments by reflexion from a calcite crystal. The reflected beam passed through a narrow cylindrical collimator 2·5 mm. in diameter and 3·6 cm. long and into the chamber through a thin aluminium window.

In the third experiment the source of X-rays was a Victor deep therapy tube worked at a peak potential of 150 kV, the total radiation from which was filtered by 4·5 mm. of copper and collimated as in the first two experiments. At first sight it might appear that the homogeneity of the beam so obtained would be very much inferior to that which could be obtained by crystal diffraction. It was found, however, that in order to obtain sufficient intensity with the power at our disposal (the maximum potential of the transformer was 150 kV) the slits had to be opened so far that the homogeneity of the reflected beam was not superior to the filtered one, and we knew rather less about the distribution of intensity in the reflected band of radiations.

The three experiments described were made with argon and measurements were confined to those electrons associated with a secondary Auger electron due to the internal conversion of K fluorescent radiation, thus ensuring that the measured photoelectrons came from the K shell. In the experiment on the lowest energy photoelectrons a 10 % argon-hydrogen mixture was used in the attempt to reduce the number of nuclear collisions near the origin of the tracks. In the other experiments the argon was diluted with nitrogen.

One of the most important factors affecting the result of a statistical experiment of this kind is the choosing of tracks for measurement. Ideally, every track appearing in the photographs which belongs to the particular class under consideration should be measured. Practically, this is impossible, since the track origins are sometimes obscured by wisps of cloud or the ends of other tracks. The procedure adopted is to measure every track that can be identified with certainty, and the aggregate of tracks chosen is assumed to be representative of the total number which actually occur. This assumption may not be justified if peculiarities in the illumination, or the geometrical arrangement of the cameras, leads to the favouring of certain directions for clear photography of the tracks. Thus tracks which start out in directions at small angles with the direction of the camera axis will rapidly pass out of focus, and will give indistinct, foreshortened images, very unfavourable for identification and for measurement. In the case where stereoscopic cameras are used (axes parallel or at a small angle with one another) any track which is unfavourable to photography in one camera will also be unfavourable to photography in the other. Thus it would be expected that the space distribution, of the tracks chosen on photographic pairs taken with stereoscopic cameras, would be deficient in tracks with directions near those of the camera axes. This deficiency will affect the longitudinal function $n(\theta)$, because $n(\theta)$ is representative of the true longitudinal distribution only if the function $J(\theta, \phi)$ can be put in the form $P(\theta) \times G(\phi)$, where $G(\phi)$ is either a constant or is a function of ϕ, independent of θ.

Williams, Nuttall and Barlow (1928) used stereoscopic cameras and found evidence that the above kind of systematic omission of tracks in the direction of the camera axes was taking place. In order to eliminate the effect on the function $n(\theta)$ they only accepted those tracks with azimuthal angles greater than 15° with the azimuthal plane containing the camera axes.

In the present experiment the cameras were mounted at right angles. This is superior to the stereoscopic method of mounting because a track which is directed towards one camera lies in the object plane of the other camera; and the indistinct photography on the first plate is accompanied by exceedingly clear photography on the other plate. Thus, with cameras mounted at right angles, it is to be expected that there will not be systematic omission from measurement of tracks at small angles with the camera axes.

The *illumination* is a very important factor influencing the choice of tracks. If it is confined to a narrow region along the X-ray beam, the number of tracks chosen, which make small angles with the axis of the beam, will be in excess. Evidence for this kind of effect was found in the first experiment when the illumination used was the older mercury-arc type; when the mirror arc was substituted, which illuminated the whole of the chamber, there was no indication of a surplus of small-angle tracks.

All tracks making small angles with the X-ray beam are favoured by the fact that they are well focused in both cameras. This is a further factor which may lead to a surplus of such tracks. However, such an excess is not revealed to any marked degree in the longitudinal distribution obtained, since the solid angle corresponding to a range $d\theta$ is comparatively small when θ is near 0 or π.

In the present experiment, all tracks are chosen after examination of the plates by two observers. and all measurements were made by these two observers in conjunction.

EXPERIMENTAL RESULTS

In order to determine the bipartition angle and the asymmetry of the distribution, a sum curve was drawn showing the experimentally observed variation with θ of the total number of photoelectrons whose directions make angles between 0 and θ with the X-ray beam. The observed longitudinal distributions are compared with those predicted theoretically by plotting points at 10° intervals which represent counts of the number of photoelectrons ejected into the 20° interval $(\theta-10°)$ to $(\theta+10°)$. The sum curves and distributions of photoelectrons are shown in figs. 1, 2 and 4, which also show the theoretical distributions by full-line curves. The use of a large interval such as 20° gives large numbers of tracks in each count with a correspondingly reduced importance of statistical fluctuations, while the effect on the theoretical curve is negligible.

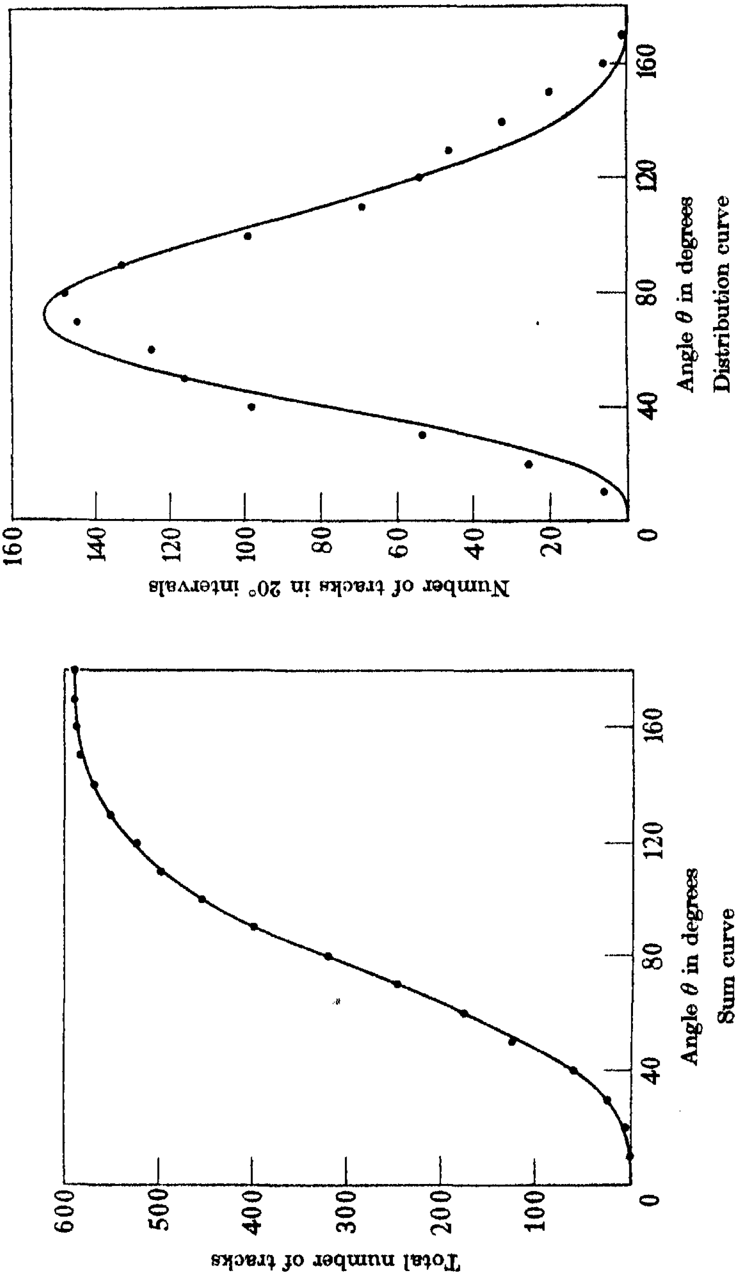
Exp. 1—Nominal wave-length 0·710 A; the slit system limited the range of admitted wave-lengths to less than 0·02 A. The value of β calculated non-relativistically is 0·236.

Exp. 2—Nominal wave-length 0·247 A; the slit system limited the range of admitted wave-lengths to less than 0·02 A. The non-relativistic value of β is 0·428.

Exp. 3—A beam of X-rays of nominal wave-length 0·135 A was obtained by filtering the total radiation from a Victor deep therapy tube through 4·5 mm. of copper. The tube was run at 150 kV (peak) and 4 mA, directly from the secondary of the transformer, a kenotron being included in the anode circuit to improve the stability and protect the tube against excessive inverse potential. The intensity-wave-length distribution curve for the total radiation was drawn by the aid of a curve published by Dauvillier (1924) for a similar tube excited at 100 kV (peak) pulsating potential, and the equation $\lambda_m = 1{\cdot}3\lambda_0 + 0{\cdot}05$ (A) which gives the relation between the short wave limit λ_0 of the general radiation and the wave-length λ_m of maximum intensity when the tube is excited by a pulsating potential (Dauvillier 1924). The intensity-wave-length distribution for the filtered beam was obtained by using values of μ/ρ for copper due to S. J. M. Allen (Compton and Allison 1935), and was corrected for the selective absorption by the argon in the chamber. It was assumed that the probability of ejection of a photoelectron from the argon atoms was proportional to λ^3.

The centre of gravity of the filtered beam corresponded to a wave-length 0·135 A. A check on this estimated wave-length was obtained by measuring the absorption coefficient of the filtered beam in copper by means of an ionization chamber. A straight line was obtained when log I_d was plotted against d the thickness of absorber for a sixfold variation in I_d. The average value of μ/ρ in copper was found to be 0·67 which corresponds to a wave-length 0·14 A (S. J. M. Allen).

The non-relativistic value of $\beta = 0{\cdot}588$. A curve was drawn to represent the theoretical distribution for the whole range of wave-lengths in the



Distribution curve

Sum curve

FIG. 1—Distribution curve for $\beta = 0{\cdot}236$.

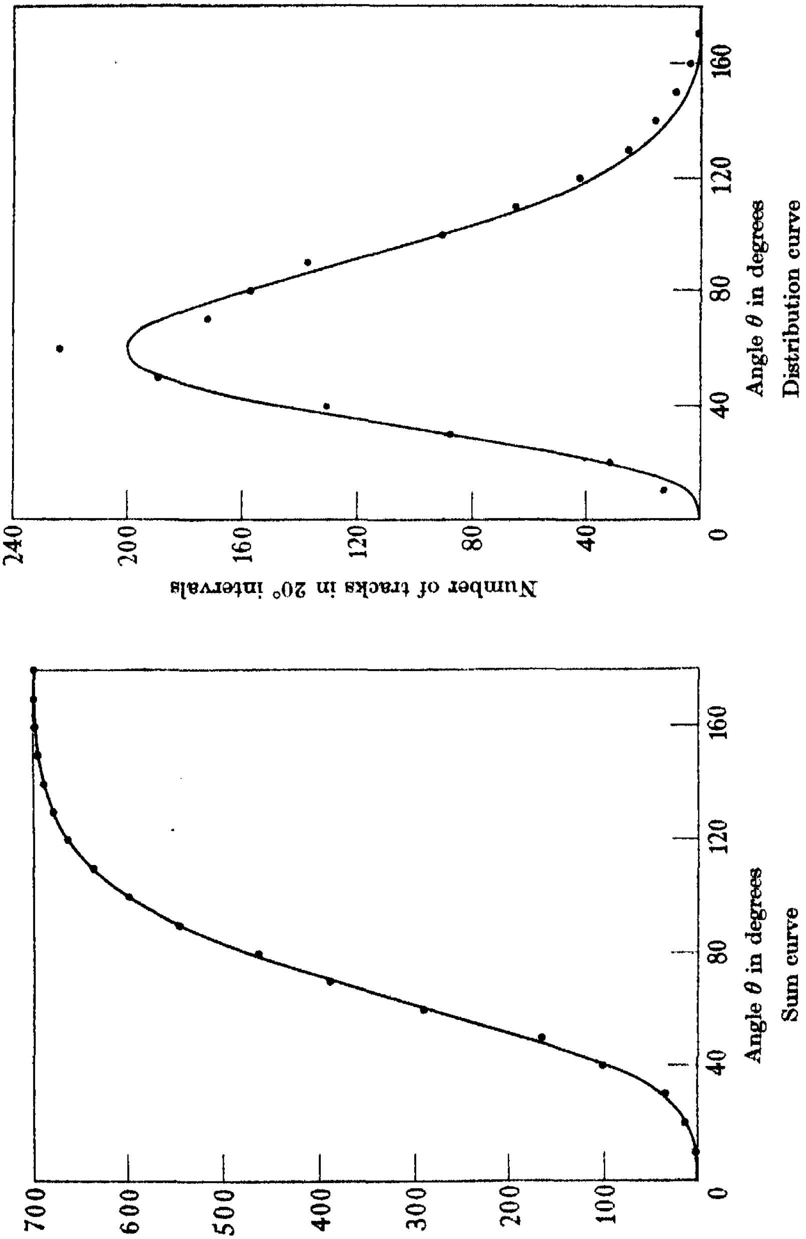


FIG. 2—Distribution curve for $\beta = 0{\cdot}428$.

filtered beam, taking into account the variation of their intensities. This curve could not be distinguished from the theoretical distribution curve for the nominal wave-length $0{\cdot}13_5$ A.

Table III gives the particulars of the three experiments, the numerical results obtained, together with the theoretical results* predicted by wave mechanics.

TABLE III—DISTRIBUTION OF K PHOTOELECTRONS FROM ARGON

λ A	$E_{\rm kin}$ e-kV	β	Tracks (total number)	θ_b		A	
				Exp.	Theory	Exp.	Theory
0·71	14·2	0·24	589	76°·8	76°·6	2·0	2·0
0·25	46·8	0·43	700	66°·8	65°·9	3·5	3·7
$0{\cdot}13_5$	88·2	0·59	412	56°·4	57°·3	6·5	6·4

DISCUSSION OF EXPERIMENTAL RESULTS

In each of the three experiments the points representing the experimental distribution counts lie closely around the theoretical curves. In the first experiment the points for angles below 50° and above 120° seem a little too high, but this was probably due to the use in this experiment of the mercury arcs. With these the only region of the chamber effectively illuminated was a region along the X-ray beam and consequently tracks making small angles with the beam were more favourable to clear photography.

Auger, during the course of his experiments, suspected the existence of two subsidiary maxima in the distribution curve. During the course of the first experiment we have also seen evidence of this effect, but further counts smoothed out the humps. It is perhaps significant that in this case our illumination was identical with that used by Auger.

It may be concluded from the observations made that the distribution curve predicted by non-relativistic wave mechanics adequately describes the experimental facts in the range of photoelectron energies from 14 to 90 kV. With regard to the values of the bipartition angle and the

---

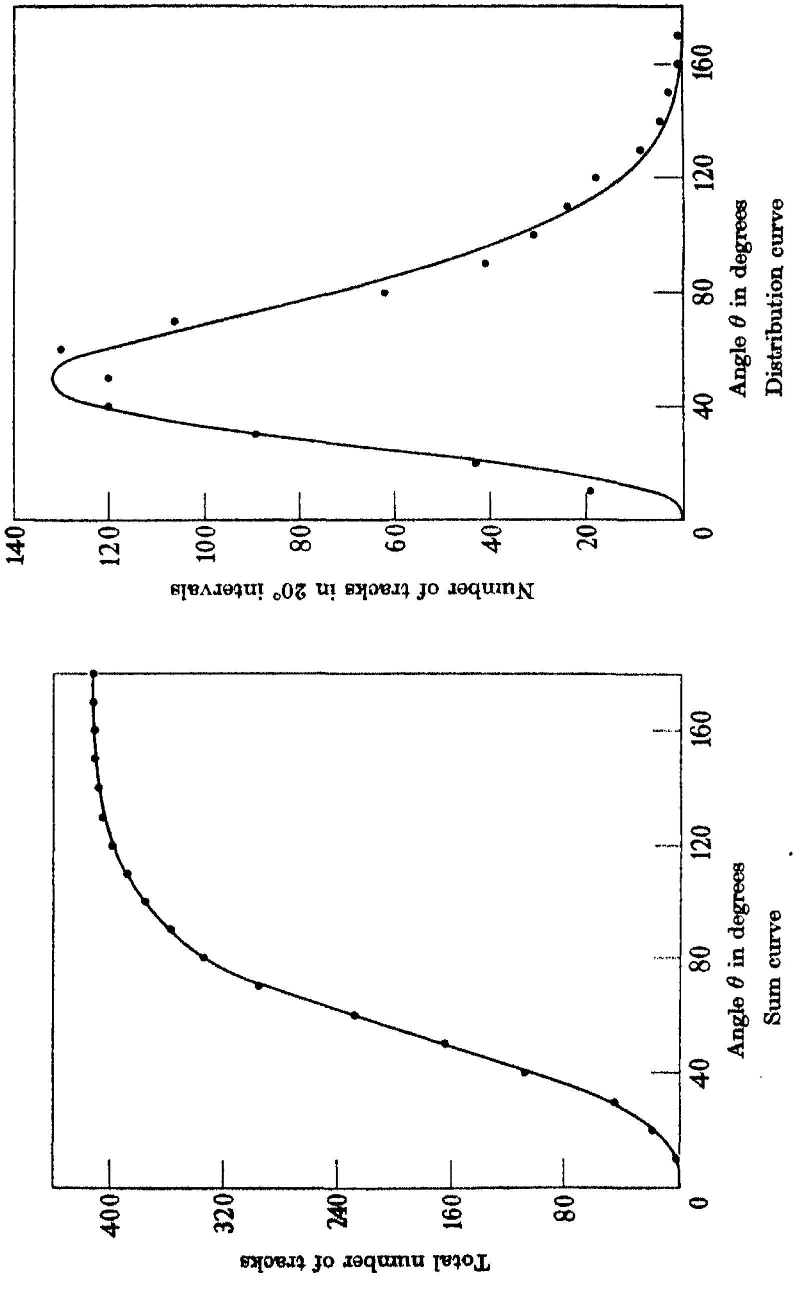
* The theoretical value of the bipartition angle is calculated from the expression

$$\cos\theta_b = \beta/(1+k),$$

and the asymmetry from the expression

$$A = \frac{(1+k+\beta)^2}{(1+k-\beta)^2}\cdot\frac{(2+k-\beta)}{(2+k+\beta)},$$

where $k = h\nu/2mc^2$.



Distribution curve

Sum curve

FIG. 3—Distribution curve for $\beta = 0{\cdot}588$.

asymmetry, the agreement between theory and experiment is closer than would be ordinarily expected from such a statistical experiment.

The value of the bipartition angle offers a very good final test of the distribution since it is less subject to statistical fluctuations than either the asymmetry or the individual counts of the distribution. In fig. 4 the variation of the bipartition angle with β predicted by wave mechanics is shown by a full line, together with the experimental values of the authors and those quoted in Table I. For the sake of interest, the values of θ_b predicted by the classical theory are indicated by a broken line. According to this theory the forward bias of the photoelectrons was caused

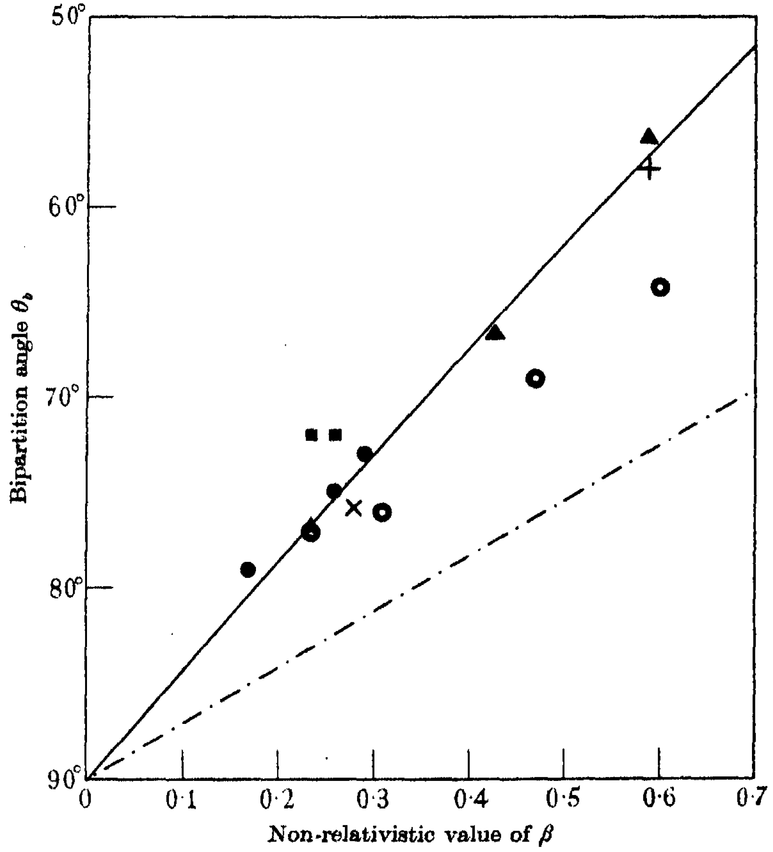


Fig. 4—Variation of θ_b with β. —— wave-mechanics theory; — · — classical theory; ▲ authors; ⊙ Auger; ● Anderson; + Lutze; × Williams, Nuttall and Barlow; ■ Loughridge.

by the electron assuming the momentum $h\nu/c$ of the incident photon resulting in the emergence of the electron at an angle θ_0, given by

$$\cos\theta_0 = \tfrac{1}{2}\beta.$$

Fig. 4 shows in a clear manner the excellent agreement which exists between the results of the authors' experiments and wave mechanics theory. We are unable to offer any satisfactory explanation for the consistent departure of Auger's results from the predictions of this theory. Our own measurements for the highest value of β (0·59) are in good agreement with the single experiment of Lutze (1931).

We wish to express our indebtedness to Professor T. H. Laby for his interest and suggestions during the course of these experiments.

SUMMARY

The longitudinal angular distribution of X-ray photoelectrons from the *K* shell of argon has been investigated with the Wilson cloud expansion chamber. A graphical method was employed in the computation of angles of ejection which combined the advantages of simplicity and accuracy.

A survey of existing data is made and the discrepancies of this with quantum theory are discussed.

The distribution curves, bipartition angles and asymmetries for photoelectron energies of 14·2, 46·8 and 88·2 e-k V have been obtained, and are found to be in excellent agreement with the predictions of wave mechanics.

REFERENCES

Anderson, C. D. 1930 *Phys. Rev.* **35**, 1139–45.
Auger, P. 1928*a* *C.R. Acad. Sci., Paris*, **187**, 1141–2.
— 1928*b* *J. Phys. Radium*, **9**, 225–30.
— 1929 *C.R. Acad. Sci., Paris*, **188**, 447–50.
Auger, P. and Meyer, T. 1931 *C.R. Acad. Sci., Paris*, **192**, 672–3.
Compton, A. H. and Allison, S. K. 1935 "X-Rays in Theory and Experiment", 2nd ed. London: Macmillan and Co.
Dauvillier, A. 1924 "La Technique des Rayons X", pp. 134 and 177. Paris: Les Presses Universitaires de France.
Fischer, J. 1931 *Ann. Phys., Lpz.*, **8**, 821–50.
Kirchner, F. 1927 *Ann. Phys., Lpz.*, **84**, 899–906.
Loughridge, D. H. 1927 *Phys. Rev.* **30**, 488–93.
Lutze, E. 1931 *Ann. Phys., Lpz.*, **9**, 853–64.

Martin, L. H. and Eggleston, F. H. 1937 *Proc. Roy. Soc.* A, **158**, 46–54.
Sauter, F. 1931 *Ann. Phys., Lpz.*, **11**, 454–88.
Schur, G. 1930 *Ann. Phys., Lpz.*, **4**, 433–42.
Sommerfeld, A. 1930 "Wave Mechanics." London: Methuen.
Watson, E. C. 1927 *Phys. Rev.* **30**, 479–87.
Williams, E. J., Nuttall, J. M. and Barlow, H. S. 1928 *Proc. Roy. Soc.* A, **121**, 611–25.
Wilson, C. T. R. 1923 *Proc. Roy. Soc.* A, **104**, 1–24.

The Absorption Spectrum of Formic Acid in the Vacuum Ultra-Violet

By W. C. Price[*] and W. M. Evans

Physical Chemistry Laboratory, Cambridge, and University College, Swansea

(*Communicated by R. G. W. Norrish, F.R.S.—Received* 14 *May* 1937)

[Plate 4]

Absorption spectra in the far ultra-violet region of the spectrum have recently assumed an important role in fixing the electronic structures of polyatomic molecules. This has been especially true of organic molecules such as acetylene, ethylene, the alkyl halides, alcohols, ethers and ketones. While all "molecular electrons" (i.e. those not contained in inner shells) can be expected to give rise to absorption bands in the region 2000–1000 A, it most frequently happens that one special electron type dominates the absorption. For example, the excitation of non-bonding $p\pi$ electrons dominate the absorption of methyl and ethyl iodides (Price 1936*a*); so-called "lone pairs" located on the oxygen atoms are responsible for all the strong bands of water, formaldehyde, etc. (Mulliken 1935*a*, *b*; Price 1935*a*, 1936*b*). In order to obtain discrete absorption bands, which are desirable for the purposes of interpreting electronic structure, it is usually necessary to take the very simplest organic molecule containing the group we wish to study. Thus for molecules of the type R_1COOR_2 (R being an alkyl group or a hydrogen atom) it has been found that only the simplest of these, namely formic acid, shows discrete absorption bands. The interpretation of the electronic structure of the carboxyl group will therefore depend to a considerable extent upon the analysis of these discrete bands. From the discussion which follows it will be easy to see why the continuous absorption from the larger molecules of the type R_1COOR_2 follows roughly the envelope of the discrete absorption of HCOOH except in so far as it is enhanced in certain regions by additional absorption from C—C and C—H bonding electrons or suffers small shifts to longer wave-lengths as a result of the substitution of hydrogen atoms by alkyl groups.

---

[*] Senior 1851 Exhibitioner, Trinity College.

The experimental technique employed in obtaining absorption spectra in the vacuum ultra-violet has been described elsewhere (Collins and Price 1934). The Lyman continuum serves as the background against which the absorption is observed, and the gas under investigation is allowed to flow continuously through the body of the spectrograph. Absorption in the region below 2000 A is usually so intense that partial pressures of a few hundredths of a mm. in a path length of the order of a metre are quite adequate to bring the bands out strongly. The spectrum of formic acid was photographed by both authors independently on two different spectrographs and with two different specimens of pure formic acid. The photographs obtained were practically identical. The first strong absorption of formic acid starts around 1550 A, at which wave-length a set of somewhat diffuse bands appear (see fig. 1, Plate 4). Between 1400 and 1100 A the bands are sharp. The analysis shows that they consist of a number of electronic transitions each of which is represented by a short vibrational pattern of two or three main members. (*Note.* The vibrational progression expected from a diatomic molecule becomes a vibrational pattern in a polyatomic molecule where many types of vibration are possible.) These patterns are strong and well separated at long wave-lengths but become weaker and overlap towards shorter wave-lengths, eventually merging into a continuum around 1100 A. A plausible Rydberg series formed from corresponding members of the patterns marked B_n (see fig. 2) is given below:

$$\nu = 91{,}370 - \frac{R}{(n-0{\cdot}40)^2} \quad (n = 2, 3, 4 \text{ etc.}). \qquad (1)$$

Actually the third strong member of the pattern was chosen to form this series; the reason being that this particular member seemed to retain its intensity better than the others in the states of higher electronic quantum number. The order of the agreement of the calculated frequencies with those observed is shown in Table I.

This agreement is considered to be sufficiently close in view of deviations likely to result from the structure of the molecule. What is most important is that the limit of the series, i.e. 91,370 cm.$^{-1}$ or 1094·5 A or 11·29 V coincides quite well with the wave-length at which the continuous absorption begins. The other features of the spectrum, such as the diminution in intensity and the crowding together of the bands as they approach the limit, tend to show that it is a real ionization continuum. There is unfortunately no experimental value of the ionization potential of HCOOH as determined by electron impact methods with which to compare this spectroscopic value. However, it is considered that the discussion to follow can establish the existence of ionization in the continuum below 1100 A with considerable certainty. The value 11·29 V obtained above for the ionization potential of formic acid may be compared with the value 10·83 V obtained for the ionization potential of formaldehyde by a similar method (Price 1935*a*). Now Mulliken (1935*b*) has shown that the latter ionization potential is to be associated with the removal of a non-bonding "$2p$" electron from the oxygen atom of the C=O group (i.e. the lone pair on the

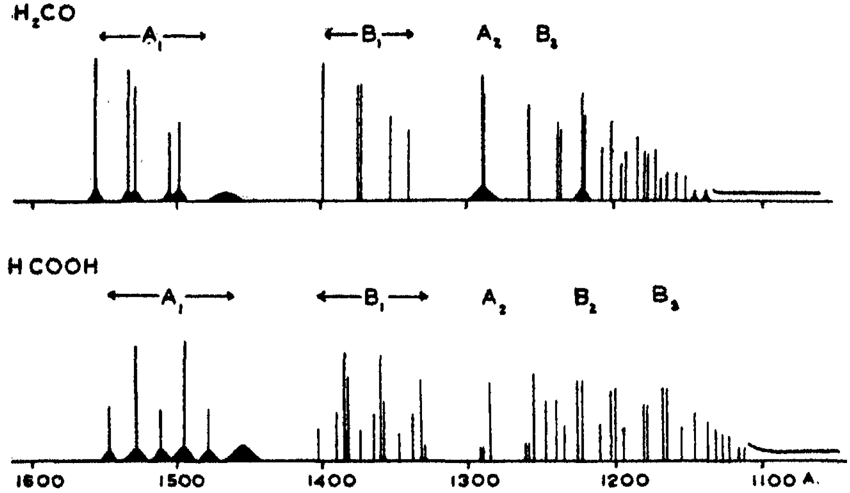


FIG. 2—Diagrams of the absorption spectra of formaldehyde and formic acid in the region below 1500 A.

TABLE I

Band	ν obs.	ν calc.	diff.
B_1	72,210	72,318	−108
B_2	81,880	81,877	+3
B_3	85,700	85,702	+2
B_4	87,670	87,607	+63
B_5	88,680	88,691	−11

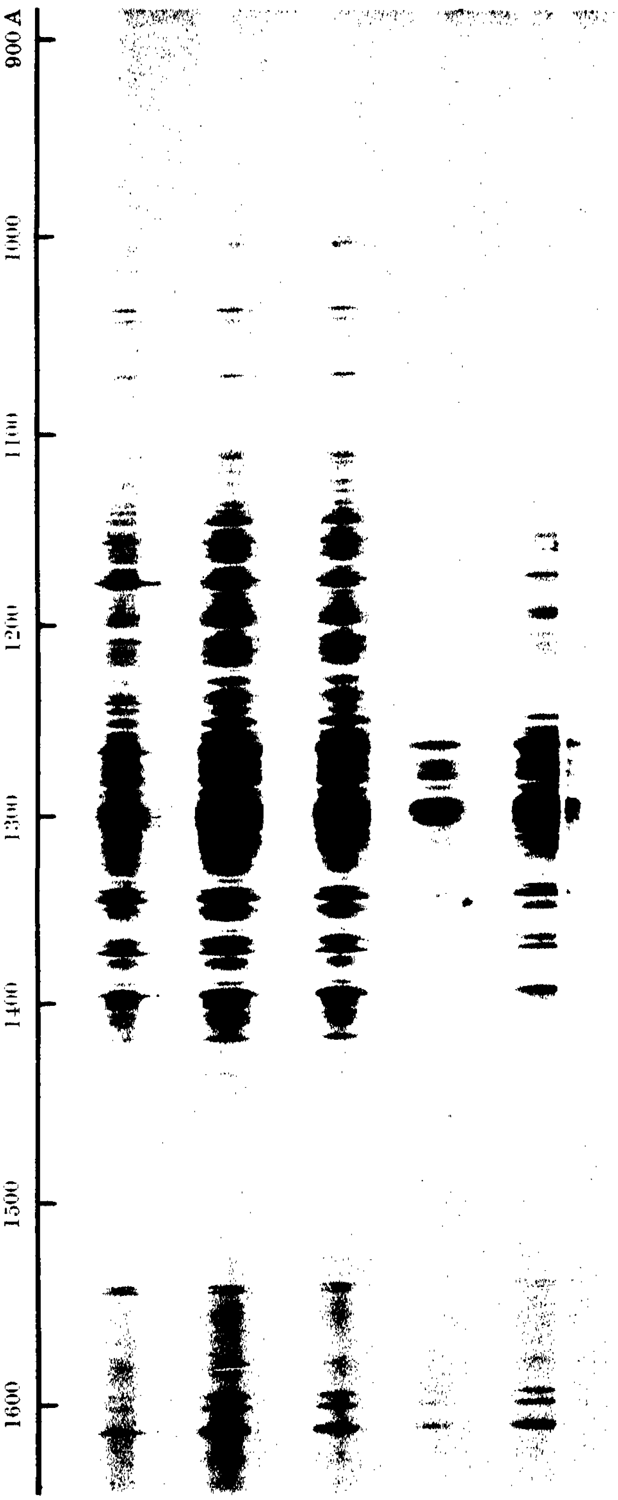


FIG. 1—The absorption spectrum of formic acid in the far ultra-violet.

oxygen). Towards the end of the paper referred to above he also states that the absorption spectra of organic acids and esters should resemble the aldehydes and ketones in that their minimum ionization potentials should correspond to the removal of non-bonding electrons from the oxygen atoms. The closeness of the two ionization potentials seems to confirm his view. In fact, below 1600 A a definite correspondence seems to exist between the electronic transitions of formaldehyde and formic acid, the related bands of both molecules occurring in the same wave-length regions. This is illustrated in fig. 2. The bands considered to belong to similar types of electronic transitions are marked by the same letter. While the vibration pattern is necessarily different for the two molecules it is clear that well-defined sets of bands occur within the same wave-length regions. This is especially conspicuous for the first few sets.

There are no *strong* bands in the spectrum of formic acid or any molecule of the type R_1COOR_2 which are analogous to the formaldehyde bands occurring in the neighbourhood of 3000 and 1745 A.* Much weaker continuous absorption does occur in these regions, but it appears that the transitions are almost completely suppressed when the carbonyl group is incorporated in a carboxyl.† Now the strong near ultra-violet absorption bands of aldehydes and ketones have usually been associated entirely with the carbonyl part of the molecule. Their suppression in the absorption spectra of molecules of the type R_1COOR_2 might be regarded as indicating that the $C=O$ group is not electronically the same in acids as it is in ketones. However, there is a considerable amount of evidence to be considered in this connexion and we shall now proceed to deal with it.

All the formic acid bands below 1500 A which seem to have counterparts in the spectrum of formaldehyde are probably of a Rydberg type, their upper states being represented by atomic [O] orbitals. The formaldehyde bands at 3000 and 1745 A are not, however, included among the Rydberg bands of that molecule. Price (1935*a*) states that "The extremely diffuse band at 1745 A can possibly be a lower member of class I though it is very far off the series." Duncan (1935), in discussing the analogous bands in acetone, states that "Originally it was thought that three Rydberg series were present in the bands and that the 51,171 cm.$^{-1}$ (1960 A) band fitted into a series...it now appears necessary to eliminate two of these series." That there is something different about the longer wave-length bands is supported by Mulliken's interpretation of the λ 3000 band of formaldehyde. He concludes rather definitely that the electron transition involved in the bands is of the type $(2p_y)_O \to [a-a]$, where $[a-a]$ is a $C \leftrightarrow O$ antibonding orbital, either $[z_{CH_2}-z_O]$ (probably) or $[x_C-x_O]$ (possibly).* The orbital $[z_{CH_2}-z_O]$ is not entirely localized in the CO bond. In addition to its C—O antibonding power it has some C—H bonding or antibonding power. While it can be shown (Mulliken 1935*b*) that it is only *slightly modified* in going from H_2CO to RHCO or R_1R_2CO, the orbital might be considerably altered if one of the hydrogens is replaced by an OH group. This seems to be a possible explanation of the suppression of the band when the carbonyl group occurs in an acid or ester. An additional factor likely to have considerable effect upon the orbitals of the CO bond is that of resonance. According to calculations from thermochemical data (Pauling and Sherman 1933), small deviations from the purely additive laws of bond energies occur when carboxyl groups are involved. This is attributed to a certain degree of resonance between the carbon-oxygen double and single bonds. While much of this energy can be accounted for by resonance which takes place in the dimer, there is still a considerable amount present in the monomeric form with which we are dealing in this work (Gillette and Sherman 1936). Such resonance would change appreciably all the orbitals of the double bond and would naturally have considerable effect upon absorption bands which have excited orbitals of this bond for their upper states. Hence it is possible that the suppression of the 3000 and the 1750 A bands in formic acid may result from changes in the bonding orbitals due to resonance between the CO bonding electrons. This resonance does not lead to complete equivalence of the CO bonds but is actually very limited in extent. Robertson and Woodward (1936) have made a careful investigation of the crystal lattice of oxalic acid dihydrate and have shown that the two oxygens are not equidistant from the carbon atom. The polymerization of the acids to form dimers and the results of the crystal analyses of acid hydrates show that the oxygen atoms are different. In addition, the persistence of the carbonyl frequency in passing from ketones to acids indicates that nothing very drastic has happened to the C=O bond (Hibben 1936). Recently, Bonner and Hofstadter (1937) have shown as a result of a study of the infra-red absorption spectrum of formic acid vapour that the force constant of the C—O bond of formic acid is slightly greater than its normal value. This also seems to show the existence of a certain limited amount of single-double bond resonance in the molecule. However, from the similarity of the spectrum of formaldehyde and formic acid below 1550 A one can say that the effect which the resonance between the bonding electrons of the carboxyl group has on the lone pairs of the oxygen atoms must be negligibly small. Only those transitions of the lone pairs which have excited orbitals of the CO bond as upper states seem to be greatly affected.

---

* For acetaldehyde the second band is around 1865 A and for acetone around 1960 A (Scheibe, Povenz and Lindstrom 1933).

† Near ultra-violet absorption measurements on some typical molecules containing ketonic and carboxyl groups are given by Henri (1919) and Ley and Arends (1932).

---

* For a complete explanation of the above notation Mulliken's original paper should be referred to (1935*b*).

VIBRATIONAL STRUCTURE

The complete analysis of the spectrum is complicated by the possible appearance of a large number of modes of vibration. The following are the Raman frequencies of formic acid given by Kohlrausch, Köppl and Pongratz (1933): 200 ± 50, 675, 706, 1059, 1201, 1396, 1675 ± 21, 1728 and 2957 cm.$^{-1}$. These values refer to the pure liquid (or solution in non-polar solvents) for which association to a dimeric form is almost complete. They are thus characteristic of $(HCOOH)_2$ in which the two carboxyl groups join together to form an eight-membered ring. The measurements of Coolidge (1928) on the vapour density show that at ordinary temperatures and pressures the association is about 98 %. Even at a pressure of 3 mm. about 50 % of the acid is present in the dimeric form. However, at a pressure of a few hundredths of a mm., which is about the pressure used in these experiments, the extrapolation of Coolidge's results show that the association is at the most about 5 %. Hence the spectra reported in this article are to be attributed almost entirely to the monomer. In this case the Raman frequencies given above have to be changed somewhat. The 200 cm.$^{-1}$ frequency certainly belongs to the dimer and need not be further considered. If it belonged to the monomer then a very large fraction of the molecules would be initially present in an excited vibrational state. For the other frequencies the Boltzmann factor is so small that all the strong bands must be attributed to transitions from a vibrationless ground state. The frequency with which we shall be mostly concerned is that at 1657 ± 21 cm.$^{-1}$. This vibration is known on well-established grounds to be the valence vibration of the C=O bond (Hibben 1936). It is found in aldehydes, ketones, acyl halides, amides, esters, acids and anhydrides and suffers only minor changes from molecule to molecule. The value 1657 ± 21 is a modified value of the frequency due to the dimer formation. In the single molecule its value is probably greater than this by about 50 cm.$^{-1}$ (Gillette and Daniels 1936). If, as has already been suggested, the excitation occurs in the carbonyl part of the molecule, then this mode of vibration with a somewhat different frequency value might be expected to appear in the vibration pattern of the excited electronic states. This is found to be the case.

In order that the differences which arise in the vibrational patterns may be easily seen, diagrams of the bands associated with the electronic transitions A_1 and B_1 have been made (fig. 3). In these the frequency differences characteristic of the pattern have been inserted.* A glance at the vibrational patterns of A_1 and B_1 shows that the only vibrational frequency which appears with more than a single quantum is that having a value between 1450 and 1500 cm.$^{-1}$. Other frequency differences occur more than once, but only as single quanta accompanying 0, 1, 2, etc. quanta of the 1450–1500 cm.$^{-1}$ vibration. This latter vibration therefore gives the periodic character to the pattern. It will subsequently be referred to as ω_1. From the existence of a number of members of the ω_1 progression it can be inferred that this frequency has been changed by a fair amount as a result of the excitation. Experience shows that this is almost certainly a small diminution. Thus ω_1 is to be associated with either the 1657 or the 1728 cm.$^{-1}$ vibration of the ground state, most probably the former. This has already been attributed to the C=O valence frequency, and thus it appears that both the excitation and the main vibration are to be associated with the carbonyl group. Regarding the other frequencies which show up very little can be said with any certainty. The values around 600 cm.$^{-1}$ probably represent vibrations perpendicular to the plane of the molecule. Those around 800 and 900 cm.$^{-1}$ may be modified CH or OH deformation,

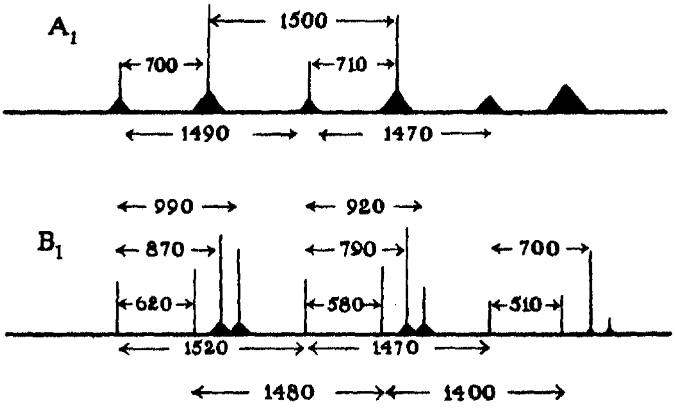


FIG. 3—Diagram showing frequency separations in the A_1 and B_1 states of formic acid.

---

* On account of the diffuseness of the bands of the A_1 state measurements on them are not so accurate as are those of the B_1 state. Actually the possible error in the measurement of the B_1 bands is about 20 cm.$^{-1}$. That of the A_1 bands may be considerably greater.

or C—O valence frequencies. Since the formic acid molecule is completely unsymmetrical, Herzberg and Teller's selection rules (1933) cannot be applied. However, the extension of the Franck-Condon principle to polyatomic molecules will still hold, and those vibrations whose controlling force fields are affected by the excitation might be expected to appear. While it has been stated here in conformity with Mulliken's interpretation of the electronic structure of the carbonyl group that the excited electron is non-bonding, this is only the case in so far as charge transfer effects have been neglected (Mulliken 1935*c*). The excitation of this electron must necessarily alter the dipole energy of the C=O bond, which in this case is probably a considerable fraction of the total bonding energy, and it is probably as a result of this effect that the C=O valence frequency appears.

Before concluding, mention should be made of the work of Scheibe, Povenz and Lindstrom (1933). These authors, using a fluorite spectrograph and the H_2 continuum, have photographed the absorption spectra of formic and acetic acids, methyl and ethyl acetate, and methyl formate down to about 1650 A. The absorption is mostly continuous and can probably be regarded as being due to the spreading to long wave-lengths of diffuse bands which are analogous to A_1. This is partly a result of increasing the pressure of the absorbing substance and partly due to the alkyl substitution which gives rise to a real electronic shift to the red on account of charge transfer effects. The limitations of the H_2 continuum prevented these experimenters from discovering the discrete bands of formic acid which occur below 1600 A. However, the rather diffuse structure they find with methyl acetate is particularly interesting. Bands of this substance occurring below 1750 A are found to have frequency separations of 1440 cm.$^{-1}$ (58,530 − 57,090) and 1470 cm.$^{-1}$ (58,840 − 57,370). These separations are obviously to be identified with the ω_1 frequencies which we have found to occur with about this value in formic acid. This strengthens our assignment of the ω_1 frequency to the C=O valence frequency. Similar examples of the persistence of a frequency difference with increase of the alkyl part of the molecule occur in ketones (Noyes 1937) and in ethylene and its higher derivatives (Carr and Stücklen 1936). It seems highly probable that the bands of methyl acetate discussed above are exactly analogous to the A_1 system of formic acid. This means that they have been shifted to longer wave-lengths by about 200 A or 0·9 e-volt. This shift is just about what might be expected to result from the substitution of two methyl groups for the hydrogens (Price 1935*a*). For example, the 1745 A band of formaldehyde is shifted to about 1970 A in acetone. Accompanying this shift there is a reduction of the ionization potential of the lone pair electrons on the oxygen atom by about 0·6 V (Noyes 1935). Thus in the same way the case of methyl acetate can probably be added to the already large list of molecules in which the ionization potentials of non-bonding electrons are reduced by charge transfer effects from alkyl groups.

A table giving the frequencies of some of the "ω_1" progressions of the *B* states is appended. The strong bands of the second column have been used for the purpose of forming an approximate Rydberg series on account of the fact that the other bands are relatively so weak as to be unobservable for higher electronic states. As the former bands are not the "ν_0" bands, the ionization potential determined from them is slightly high (probably by about 0·1 V). The values of the frequencies given in the table are correct to within 20 or 30 cm.$^{-1}$. The intensities are eye estimates. In addition to the bands occurring in the region 1550–1100 A, some very diffuse (pre-ionized) bands were found at shorter wave-lengths superimposed on the ionization continuum. There are also a group of bands around 1250 A which are different in character from the surrounding bands and do not fit into the analysis. These, together with some bands around 1120 A, may be due to the excitation of the lone pair from the hydroxyl oxygen. However, it has not been possible to do much with them. It should be stressed that the conclusions arrived at in this article are in no way to be interpreted as a complete analysis of the spectrum (similar, for instance,

TABLE II—TABLE OF "ω_1" PROGRESSIONS IN THE *B* STATES

State	I	ν obs.	diff.	I	ν obs.	diff.	I	ν obs.	diff.
B_1	9	71,960		10	72,210		10	72,330	
			1480			1440			1440
	9	73,440		10	73,650		7	73,770	
			1400			1440			
	8	74,840		7	75,090			—	
			1420			1380			
	4	76,260		4	76,470			—	
B_2	8	81,680		8	81,880			—	
			1400			1440			
	6	83,080		7	83,320			—	
			1380			1410			
	3	84,460		5	84,730			—	
B_3	4	85,460		6	85,700			—	
			1490			1470			
	3	86,950		5	87,170			—	
B_4		—		4	87,670			—	
						1480			
		—		2	89,150			—	

to the analysis which has been accomplished for the bands of acetylene (Price 1935*b*)). This is too much to expect in the face of difficulties such as the lack of symmetry of the molecule, its more complicated vibrational pattern and the possibility that some of the weaker bands are to be attributed to double molecules or to excitation from the hydroxyl oxygen. While a number of bands are as yet unassigned it is, nevertheless, considered that the general features of the spectrum are satisfactorily accounted for by the hypothesis that the main bands are of a Rydberg type (their upper states being probably represented by atomic [O] orbitals), and that they lead to the ionization of the lone pair electrons of the carbonyl oxygen atom.

In conclusion we wish to thank Professor J. Lennard-Jones and Dr. G. B. B. M. Sutherland for much helpful discussion in connexion with the material of this article. Also we wish to express our gratitude to the Government Grant Committee of the Royal Society for a grant for apparatus.

SUMMARY

Formic acid has been found to possess a large number of discrete absorption bands in the region below 1550 A. These consist of numerous electronic transitions each of which is represented by a short vibrational pattern. These patterns are strong and well separated at long wave-lengths but become weaker and overlap towards shorter wave-lengths, eventually merging into an ionization continuum around 1095 A or 11·3 e-volts. Reasons are given for the belief that the excitation and removal of the lone pair electrons from the carbonyl oxygen atom is responsible for the greater part of this absorption spectrum.

REFERENCES

Bonner and Hofstadter 1937 *Bull. Amer. Phys. Soc.* A, **99**, 25 (Washington Meeting).
Carr and Stucklen 1936 *J. Chem. Phys.* **4**, 760–8.
Collins and Price 1934 *Rev. Sci. Instrum.* **5**, 423–5.
Coolidge, A. S. 1928 *J. Amer. Chem. Soc.* **50**, 2166–75.
Duncan, A. B. F. 1935 *J. Chem. Phys.* **3**, 256.
Gillette and Sherman 1936 *J. Amer. Chem. Soc.* **58**, 1135–9.
Gillette and Daniels, F. 1936 *J. Amer. Chem. Soc.* **58**, 1139–42.
Henri, V. 1919 "Études de Photochimie." Paris: Gauthier-Villars et Cie.
Herzberg and Teller 1933 *Z. phys. Chem.* B, **21**, 410.
Hibben, J. 1936 *Chem. Rev.* **18**, 1–232.
Kohlrausch, Köppl and Pongratz 1933 *Z. phys. Chem.* B, **21**, 242–55.

Ley and Arends 1932 *Z. phys. Chem.* B, **17**, 177–219.
Mulliken, R. S. 1935*a* *J. Chem. Phys.* **3**, 506–14.
— 1935*b* *J. Chem. Phys.* **3**, 565–73.
— 1935*c* *J. Chem. Phys.* **3**, 514–17.
Noyes, W. Albert 1935 *J. Chem. Phys.* **3**, 430–2.
— 1937 *J. Phys. Chem.* **41**, 81–9.
Pauling and Sherman 1933 *J. Chem. Phys.* **1**, 606–17.
Price, W. C. 1935*a* *J. Chem. Phys.* **3**, 256–9.
— 1935*b* *Phys. Rev.* **47**, 444–52.
— 1936*a* *J. Chem. Phys.* **4**, 539–51.
— 1936*b* *J. Chem. Phys.* **4**, 147–53.
Robertson, J. M. and Woodward, I. 1936 *J. Chem. Soc.* pp. 1817–24.
Scheibe, Povenz and Lindstrom 1933 *Z. phys. Chem.* B, **20**, 283.

The Influence of the Difference of Orientation of Two Crystals on the Mechanical Effect of their Boundary

By Bruce Chalmers, B.Sc., Ph.D.

(*Communicated by E. N. da C. Andrade, F.R.S.—Received* 28 *May* 1937)

1—Introduction

The influence of the crystal boundaries on the mechanical properties of metals has long been recognized, and various theories have been advanced regarding the structure of the boundary region between adjacent crystals. Since in general the lattices of two neighbouring crystals growing from independently formed nuclei will not register, there must be some modification of the lattices where they join. The two principal theories as to the nature of this modification are (*a*) that a structureless or amorphous layer of a thickness of at least some tens of atoms occupies the space between crystals, and (*b*) that the atoms are arranged on a transitional lattice joining one crystal to another.* There are two ways of distinguishing between these theories. In the first place, the amorphous theory regards the atoms of the boundary layer as being disposed at random, whereas the transitional theory considers each atom as having a calculable position; and secondly, the structure and properties of the amorphous layer should be independent of the relative orientations of the axes of the bounding crystals, while a transitional lattice should vary in its structure and properties according to the angle through which the transition is made.

---

* A review and bibliography of the theories of the crystal boundary has recently been compiled by Bucknall (1936) which makes further references here unnecessary.

That the second of these criteria can be applied has already been indicated by the author (Chalmers 1937). A number of specimens of tin of cylindrical form, some consisting of a few crystals with longitudinal boundaries, were subjected to tensile tests, and the results showed that the different behaviour of different specimens could only be accounted for by the hypothesis that the effect of a crystal boundary depends on the difference of orientation of the crystals.

In the experiments to be described cylindrical specimens consisting of two crystals, separated by a longitudinal boundary, were subjected to a simple tensile test. The crystals were orientated similarly with respect to the axis of the cylindrical specimen, so that rotation through an angle A about this axis of the specimen would bring one into coincidence with the other. This ensures that the resolved shear stresses on all equivalent planes, and hence necessarily on equivalent glide planes, were the same with a given load in all experiments. Everything thus being fixed except the angle A, which varied from specimen to specimen, the variation of critical stress with A was investigated, the meaning attached to critical stress being defined in the paper. The results are discussed in terms of the boundary structure.

2—EXPERIMENTAL TECHNIQUE

The experimental work consisted in (*a*) preparing cylindrical specimens of tin, each specimen consisting of two crystals with a longitudinal boundary, (*b*) determining the orientations of the two crystals of each specimen, and (*c*) applying a tensile test to determine the characteristic stress for each specimen.

(*a*) *Preparation of "Bicrystal" Specimens*

The specimens were all made from the same ingot of "Chempur" tin (analysis, tin 99·98 %, copper 0·00132 %, antimony 0·00118 %, lead 0·00585 %, iron 0·00055 %, bismuth 0·00352 %, arsenic 0·00005 %, nickel 0·00003 %, silver 0·00018 %, zinc, cobalt and sulphur, nil) by the following method.

Glass tubes of internal diameter 3·5 mm. were filled with tin for a length of 10–15 cm. by suction from a crucible containing molten tin, and the specimens prepared by a modification of the "moving furnace" method (Andrade and Roscoe 1937) in which the specimen tube is fixed horizontally and a heating coil is moved slowly along, melting the specimen locally so that it refreezes progressively from one end. The introduction of two seed crystals of smaller diameter and of the required orientations into the tube at the end from which crystallization starts, so that the ends of the seeds are melted when in contact with the molten end of the specimen, causes crystallization to commence with two orientations, the lattices meeting at a boundary which is roughly a diameter of the cross-section. If the rates of crystal growth of both parts are equal, then the crystals extend along the whole specimen in such a way that the boundary between them is a plane containing the axis of the cylinder.

To ensure that the crystal lattices are all similarly orientated with respect to the axis of the specimen, not only for the two crystals comprising one specimen, but from specimen to specimen, the following orientation of all the seeds was adopted. The 001 axis was arranged to be perpendicular, and the 101 axis to be at 45°, to the axis of the specimen. This symmetry about the axis ensures that the resolved shear stress shall be the same on corresponding planes in different cases, but leaves the angle A between the 001 axis of the two crystals constituting a given specimen under control.

The angle A was adjusted to any desired value by manipulating the seed crystals so that their 001 axes were at the required angle before they were brought into contact with the molten meniscus from which crystallization of the specimen started.

When the orientations of the two seeds with respect to the specimen axis are different, then in general the boundary is not a diametrical plane but becomes a plane inclined to the specimen axis. This is receiving further investigation.

A single seed instead of two was used when it was required to prepare a single crystal of the same orientation.

(b) *Measurement of Orientation*

The optical reflexion method previously described (Chalmers 1935) was used for the determination of the orientations. The specimen, after etching with ferric chloride solution, was mounted on the axis of a goniometer, and a beam of slightly converging light directed radially on to it. The reflexion spots were viewed on a screen, and the positions of the goniometer circle were observed at which the 001 spot of each crystal was at a given position on the screen. The difference of the two angles gave the angle between the 001 axes of the two crystals. The identification of the 001 spot was carried out by using the method detailed in the paper to which reference has been made (Chalmers 1935).

(c) *Tensile Test*

By means of preliminary experiments it was found that a very simple tensile test would serve to determine a characteristic property of each specimen. An extensometer was constructed in which a load could be applied in units of 50 g. with a leverage of 5, and for a gauge length of 2 cm. the extension was measured with a magnification of about 250 by means of an optical lever.

As the load is increased from zero, no extension is observed until the stress reaches a value characteristic of the specimen, after which the extension continues rapidly if the stress is further increased. In order to estimate rather more accurately the minimum stress that causes a measurable extension, a definite value (0·2 mm. on the scale) was chosen as defining the minimum observable extension, and a curve was plotted of extension and load, the load being applied at a fixed rate. A typical curve is shown in fig. 1. Interpolation between *A* and *B* indicates the load that would produce the extension indicated by *CD*. No claim for great accuracy is made for this method of determining the standard stress, but it is probably accurate to ± 20 g., which is shown by the results to be adequate for the present purpose. Since the specimens were all prepared in glass tubes of the same size, the stress can be taken as proportional to the load applied.

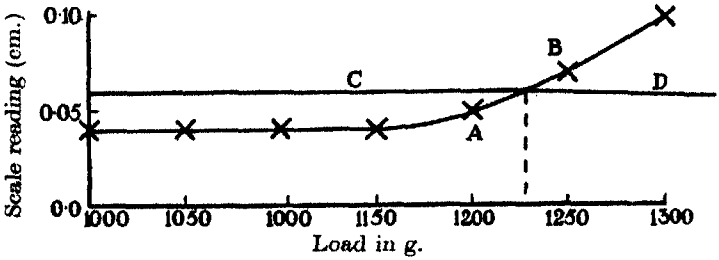


FIG. 1

Reference to the results obtained with a precision extensometer on a similar specimen prepared accidentally in an investigation on single crystals shows that microplasticity does not occur with such specimens (Chalmers 1936), and that the flow increases very rapidly when the critical stress has been exceeded; hence the values obtained here are slightly greater than, but related to, the minimum stresses that cause flow to occur.

3—EXPERIMENTAL RESULTS

The experimental results can best be exhibited in the form of a graph, fig. 2, relating the angle *A* between the 001 axes of the two crystals, with the load required to produce the minimum observable extension. It will be observed that the points fall roughly on a straight line and the point for zero angle (i.e. a single crystal) is on the same line.

Further, it has been found that the direction of the boundary plane with respect to the axes has no influence on the mechanical property under consideration, different specimens with the same angle *A* but different positions of the boundary giving the same results within the limits imposed by the experimental method.

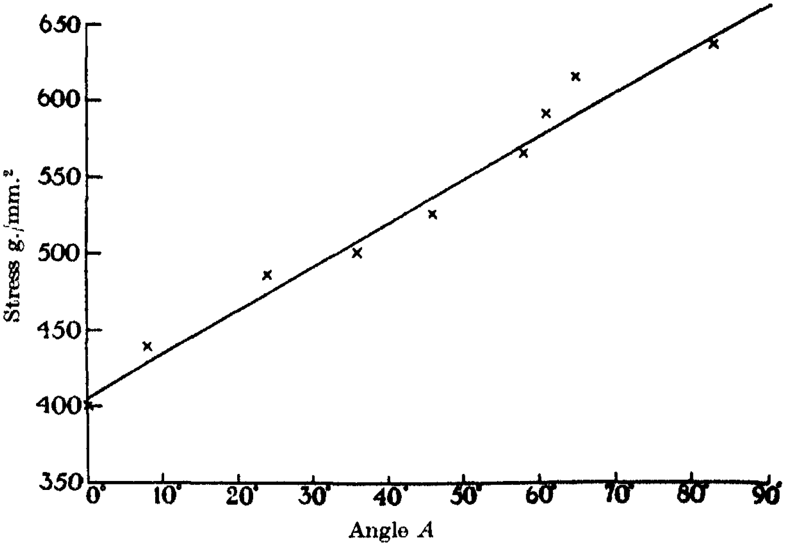


FIG. 2

4—DISCUSSION

Before discussing the implications of the results described above it may be useful to consider certain features of the experimental procedure. It will have been observed that the crystal boundary concerned is always longitudinal, and therefore parallel to the direction of application of stress. The boundary is effective over the whole gauge length but is not solely responsible for any part of the extension. The effect of the boundary may be (*a*) to introduce a thin layer having mechanical properties of its own, differing from those of the crystals and (*b*) to modify the stress-strain properties of the crystals.

This method of preparation of the specimens, i.e. direction of crystallization parallel to the boundary, should ensure that the boundary resembles that formed in the usual way in a perfectly pure metal, i.e. should be free from the impurities segregated by eutectic solidification. Application of mercury shows that preferential diffusion takes place along this kind of boundary as it does along ordinary ones (Chalmers 1937, p. 298), showing that the diffusion of mercury does not depend on the presence of eutectic impurities. Another feature of the experimental technique is that every crystal was similarly oriented with respect to the stress, and so under a given stress all the specimens would stretch by the same amount if it were not for the effect of the boundaries. The experiments deal, therefore, with the effect of the boundary in modifying the stress-strain relations of the crystals, and with the mechanical properties of the boundary layer itself. The results show that the critical stress as defined above varies from the value for a single crystal for $A = 0°$ to a maximum for $A = 90°$, the greatest possible angle.

The effect of the boundary, considered as super-imposed on a single crystal, varies from zero at $A = 0°$ to a maximum for $A = 90°$, the variation being approximately linear. On account of the lack of precision in the determination of the critical load, however, too much significance must not be attached to the actual form of the curve obtained in fig. 2. A linear relationship between load and angle is unlikely to be a true representation of the effect, and it is hoped that more refined measurements will reveal the function of the angle with which the load varies linearly.

It follows that if the boundary structure is independent of A, as the amorphous theory requires, the boundary itself makes no contribution to the mechanical property under test, and the change is entirely due to the effect of the two lattices on each other, tending to prevent extension for a large angle A more than for a small angle. Since the two crystals should stretch equally, there is no reason why each should tend to prevent the other extending.

The possible lateral distortion at the boundary that may accompany glide would vary according to the direction of the boundary relatively to the crystal axes; and since the direction of the boundary has been found to be of no significance it follows that the mutual prevention of this lateral distortion cannot be the cause of the change of mechanical properties with the angle A. In any case, if the boundary were symmetrically disposed between the axes of the two crystals the lateral distortion should be the same for both crystal faces at the boundary, which should then not hinder glide for large angles A more than for small ones. The same criticism also applies to the adsorbed layer theory, which is therefore also rendered untenable.

On the other hand, if the boundary is regarded as a transition zone between one lattice and the other, the hindrance to glide of each lattice may be controlled by the amount of transition between the two lattices, a violent transition (A large) being more effective in preventing glide than a slight transition (A small), which is in agreement with the experimental results.

A modification of the transitional lattice theory (Hargreaves and Hills 1929) suggests that two lattices may share some atoms if the inclination of the lattices happens to have certain values deduced from geometrical considerations; such a theory is not supported by the present experiments, as it should give small stresses for certain intermediate angles between 0° and 90°, which were not observed.

It would seem that the result obtained for tin is probably of general application, since the behaviour of tin under stress is very similar to that of other metals at temperatures similarly related to the melting-point, so the conclusion is that the effect of crystal boundaries in determining the mechanical properties of a crystal aggregate depends entirely on the angles between the axes of neighbouring crystals, and that this constitutes strong evidence for the transitional lattice boundary theory.

In conclusion, the author wishes to express his indebtedness to Professor E. N. da C. Andrade, F.R.S., for the encouraging interest he has taken in this work; to the International Tin Research and Development Council for a grant and to its Director, Mr. D. J. Macnaughtan, F.Inst.P., for his personal interest; and to the Governors and Principal of the Sir John Cass Technical Institute for the provision of facilities.

SUMMARY

The effect of the boundary between the crystallites in a metal specimen on the mechanical behaviour is of fundamental importance for the understanding of the strength of polycrystalline metals. The investigation deals with a particularly simple case, in which the effect of a single boundary can be measured. The specimens tested consist of two crystals with a longitudinal boundary prepared so that the orientation of the lattices of the two crystals with respect to the direction of the stress is the same in all cases, while the angle between the two lattices varies from specimen to specimen. The results show that the critical tension varies regularly with the angle between the two crystals, being a minimum when the two lattices are similarly orientated (i.e. for single crystals) and a maximum when the two

lattices are at right angles. It is concluded that the boundary has no inherent strength, and that the results are best explained as the effect of a transitional lattice. The evidence is against the existence of an "amorphous layer" or "inter-crystalline cement".

REFERENCES

Andrade, E. N. da C. and Roscoe, R. 1937 *Proc. Phys. Soc.* **49**, 152–77.
Bucknall, E. H. 1936 "Metal Industry", pp. 311–16, 369–73, and 396–9.
Chalmers, B. 1935 *Proc. Phys. Soc.* **47**, 733–45.
— 1936 *Proc. Roy. Soc.* A, **156**, 427–43.
— 1937 *J. Inst. Met.* **61** (in the Press).
Hargreaves, F. and Hills, R. J. 1929 *J. Inst. Met.* **41**, 257–83.

Scattering of Slow Neutrons

BY M. GOLDHABER, PH.D., *Magdalene College, Cambridge*, AND
G. H. BRIGGS, PH.D., *University of Sydney*

(*Communicated by Lord Rutherford, O.M., F.R.S.—*
Received 9 *June* 1937)

INTRODUCTION

The absorption and scattering of slow neutrons have been studied by various methods. In their first survey, Amaldi, D'Agostino, Fermi, Pontecorvo, Rasetti and Segrè (1935) investigated the absorption of slow neutrons by different elements inside a paraffin block. The number of slow neutrons present was determined by the radioactivity produced by neutron capture in an indicator (e.g. silver). The values for the absorption coefficients which they obtained with this arrangement can be regarded as a measure of the "true" absorption of slow neutrons. Later, Dunning, Pegram, Fink and Mitchell (1935) measured the "total" cross-sections, i.e. the sum of the cross-sections for true absorption and for elastic scattering,* using a fairly well-defined beam of slow neutrons, and a lithium-coated ionization chamber as indicator. Recently, Griffiths and Szilard (1937) have determined the cross-sections of some strongly absorbing elements using the captured γ-rays from cadmium as indicator.

---

* For brevity we shall refer to the process of "scattering plus true absorption" simply as "absorption" when no ambiguity can arise.

The scattering of slow neutrons was studied for some elements by Mitchell and Murphy (1935), Mitchell, Murphy and Whitaker (1936), Budnitzky and Kurtschatow (1935) Pontecorvo and Wick (1936), and others. In these experiments, the slow neutrons issuing from a paraffin block passed through an indicator (e.g. silver), and were scattered backwards from the substance under investigation. When thin scatterers are used, the percentage increase of the radioactivity produced in the indicator gives a measure of the scattering cross-section. This method has the disadvantage, even in the case of good scatterers, that the increase in the radioactivity of the indicator is usually small compared with the effect due to the primary neutrons. A quantitative interpretation of the results may be further complicated by the fact that the neutrons leave the surface of the paraffin block at all angles from 0 to 90°. Also, multiple scattering is not avoided.

In part II we shall describe measurements of the scattering cross-sections for slow neutrons by a method free from these objections. It consists in the determination of the cross-sections relative to that of carbon. Carbon was chosen as the standard element because its cross-section for true absorption is known to be negligible compared with its cross-section for scattering. We may therefore calculate the absolute scattering cross-sections from the relative values by multiplying by the total cross-section of carbon. To determine this we used an arrangement somewhat different from that of Dunning and others (1935) (see part I) with which we also measured the total cross-sections of hydrogen and of several other elements.

I—DETERMINATION OF THE TOTAL CROSS-SECTION

1—*Experimental Arrangement*

The arrangement used for measuring total cross-sections is shown in fig. 1.

The neutron source consisted of 100 mg. radium mixed with finely powdered beryllium (equivalent in neutron intensity to about 60 millicuries Rn + Be). The canalization method of Szilard, Hopwood and Chalmers (cf. Hopwood and Chalmers 1935) was used to obtain a beam of slow neutrons comparatively free from fast neutrons. The beam was defined by a cylindrical shell of powdered boron oxide (*B*) and a cadmium diaphragm (at *A*). The slow neutrons were detected by means of an ionization chamber filled with boron trifluoride at a pressure of two atmospheres. The chamber was shielded by boron oxide and cadmium from slow neutrons scattered from the walls of the room. The "kicks" were recorded with a linear amplifier and a scale-of-two thyratron counter. The natural effect of the ionization chamber was 0·6 kicks/min. The γ-rays from the source were absorbed by a lead cone.

The absorbers were interposed at *A*. Without an absorber, the number of boron disintegrations recorded was 150/min.; when all C neutrons were absorbed by a cadmium sheet, 1 mm. thick, only 8 kicks/min. were recorded. Of these, about 6 kicks/min. were due to neutrons of "medium" velocity which could be absorbed or scattered away by thick layers of boron oxide, paraffin, nickel, etc., interposed at *A*. We shall call these residual neutrons of medium velocity "*R* neutrons". When the *R* neutrons were removed there remained a background effect of about 2 kicks/min. due to fast neutrons from the source and to neutrons scattered from the walls of the room.

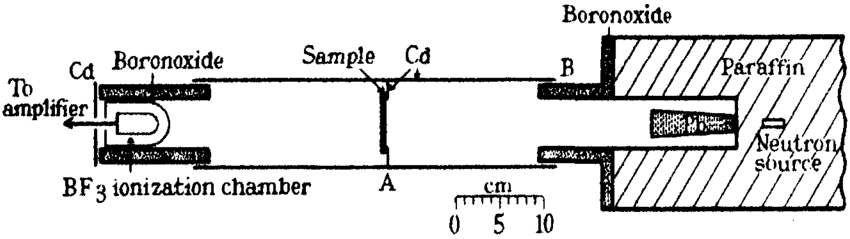


FIG. 1—Arrangement for measuring total cross-sections.

Our neutron beam, then, can be regarded practically as a beam of C neutrons. Appreciable corrections for the absorption of the *R* neutrons have to be made only for large thicknesses of the absorbers. Obliquity of the neutron beam and scattering of neutrons from the sample into the BF_3 counter will give rise to corrections which would only be of the order of 1 % and of opposite sign, and hence they will be neglected.

2—*Results*

With the experimental arrangement described in the preceding paragraph we have determined the total cross-section for the absorption and scattering of C neutrons for several elements. In figs. 2 and 3, the ratio I/I_0 of the transmitted to the initial intensity of the C neutrons is plotted on a logarithmic scale as a function of the thickness of the absorbers. The absorption follows practically an exponential law.

From fig. 2 we can deduce the mean free path λ in paraffin for thermal neutrons. We find $\lambda = 2{\cdot}5$ mm. (corresponding to 0·225 g./cm.^{2}). Amaldi and Fermi (1936) find $\lambda = 3$ mm.

The total cross-section N is defined by $\sigma_{\text{total}} = 1/N$, where N is the number of atoms per cm.^{2} of an absorber which reduces the intensity of the neutron beam to $1/e$. In Table I are reproduced the values of σ_{total} deduced from figs. 2 and 3, as well as the values for some other elements which were obtained by determining the absorption for a single thickness of the absorbers. The value for hydrogen has been calculated from the mean free path of C neutrons in paraffin, assuming an average value of $n = 30$ in the chemical formula for paraffin (C_nH_{2n+2}), and allowing for the scattering

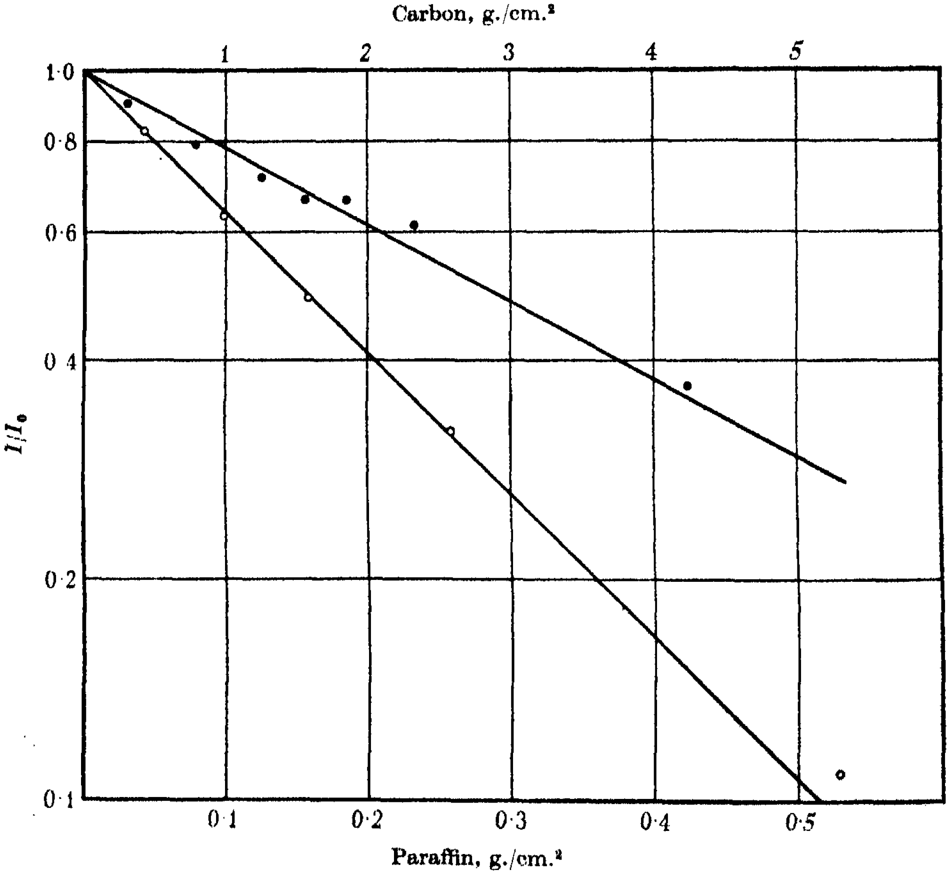


FIG. 2—Absorption curves for paraffin ○, and carbon ●.

due to carbon. In column 3 the values of Dunning and others (1935) are given for comparison.

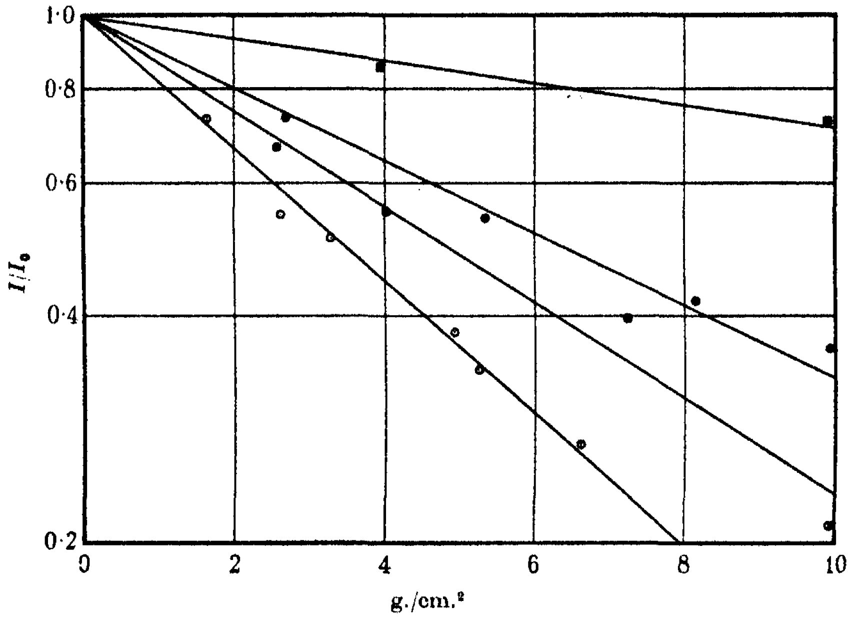


FIG. 3—Absorption curves for iron ⊗, nickel ⊙, copper ●, and lead ■.

TABLE I—$\sigma_{\text{total}} \times 10^{24}$ cm.$^{-2}$

Element		Dunning and others
1 H	47·5	35 (38*)
6 C	$4{\cdot}8_3$	4·1
15 P	13·6	14·7
26 Fe	13·6	12·0
28 Ni	19·7	15·4
29 Cu	11·9	7·5
34 Se	20·4	19
47 Ag	63	55
78 Pt	21·5	25
82 Pb	12·5	8·6
83 Bi	8·9	8·2

For some elements, our values for σ_{total} are appreciably higher than those of Dunning and others. For those elements in which the scattering cross-section is not negligible compared with the total cross-section, we might expect to obtain higher values than these authors, because in their arrangement an appreciable fraction of the neutrons scattered by the sample pass through the neutron detector. We shall return to this point later (part II, § 4).

---

* See Fink (1936).

II—Scattering of Slow Neutrons

1—*Experimental Arrangement*

To investigate the scattering of slow neutrons we have compared the numbers of slow neutrons scattered by various substances at a mean angle of about 90°.

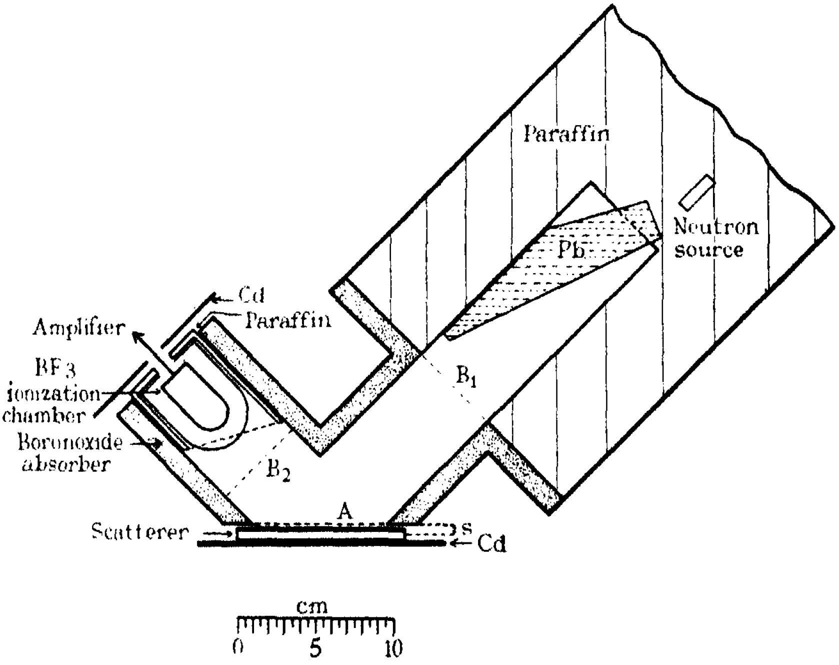


Fig. 4—Arrangement for measuring scattering cross-sections.

The apparatus is shown diagrammatically in fig. 4. The primary and scattered neutron beams are defined by mutually perpendicular cylindrical shells of boron oxide. The slow neutrons are incident at an average angle of about 45° on a sheet of the scattering element which covers the elliptical opening at *A*. Neutrons which have been scattered at an average angle of about 90° enter the BF_3 counter. To increase the number of neutrons recorded, the back and sides of the BF_3 counter were covered with paraffin wax about 2·5 mm. thick. This paraffin layer "reflects" an appreciable number of C neutrons back into the ionization chamber, thus increasing the efficiency of slow neutron detection by about 50 %.

The background effect, due mainly to fast neutrons from the source, had to be subtracted from all measurements. It is defined by the number of counts obtained with a Cd sheet at A. It is known that Cd absorbs but does not scatter C neutrons, and this is demonstrated by the fact that the number of background counts does not decrease appreciably when the primary or secondary beam is cut off by introducing a Cd sheet at B_1 or B_2. The background effect of about 23 counts/min. is small compared with the number of counts observed with a thick paraffin scatterer at A (about 300/min.).

2—*Analysis of Scattered Neutron Beam—Saturation Curves*

To analyse the secondary or scattered neutron beam into its C and R components and to study their relation with the C and R components of the primary beam we used Cd disks, 1 mm. thick, which could be interposed at B_1 and B_2.

We represent the secondary beam symbolically by $C+R$, where C and R are the number of counts per min. due to the C and R components respectively. Let C_C represent the number of counts per min. due to scattered C neutrons, C_R the number of counts per min. due to neutrons which in the process of scattering are transformed from R to C neutrons; and let R_C and R_R be defined similarly. Then by definition

$$C+R = C_C + C_R + R_C + R_R.$$

For a given scatterer at A we can determine the quantities C_C, C_R, R_C and R_R by the following four measurements:

(1) With no Cd disk inserted at either B_1 or B_2 we measure the quantity

$$I_1 = C_C + C_R + R_C + R_R.$$

(2) With a Cd disk at B_1 we measure the quantity

$$I_2 = C_R + R_R.$$

(3) With a Cd disk at B_2 we measure the quantity

$$I_3 = R_C + R_R.$$

(4) With a Cd disk at B_1 as well as B_2 we measure the quantity

$$I_4 = R_R.$$

From the values I_1, I_2, I_3 and I_4 we obtain

$$C_C = I_1 - I_2 - I_3 + I_4,$$
$$C_R = I_2 - I_4,$$
$$R_C = I_3 - I_4,$$
$$R_R = I_4.$$

For a thick paraffin scatterer (4·8 g./cm.^{2}) we find no appreciable difference between I_3 and I_4. Hence $R_C \approx 0$. On the other hand, we find $C_R = 35$. We may therefore conclude that R neutrons are slowed down in the process of scattering and transformed into C neutrons, whereas, no appreciable number of C neutrons are transformed into R neutrons. With the assumption of thermal equilibrium this indicates that Cd is not appreciably transparent to slow neutrons of energies for which the Maxwell distribution curve at room temperature is appreciably different from zero.

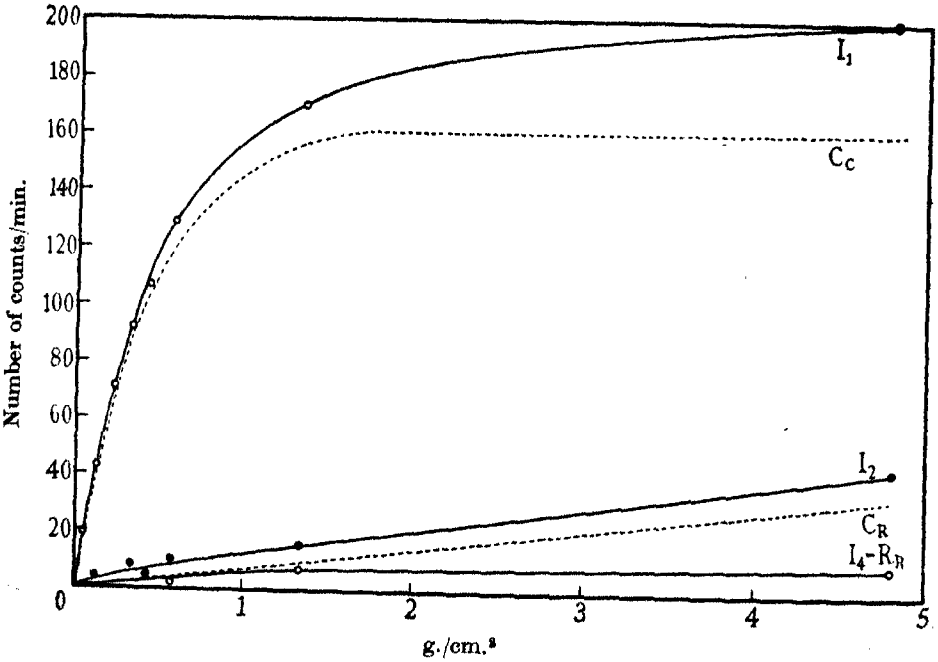


Fig. 5—Analysis of scattering curves obtained with paraffin wax.

In fig. 5 I_1, I_2 and I_4 ($= R_R$) also C_C and C_R are represented as a function of thickness for a paraffin scatterer. In figs. 6 and 7 I_1 is given as a function of thickness for nickel, copper, iron and lead. The curves of figs. 5, 6 and 7

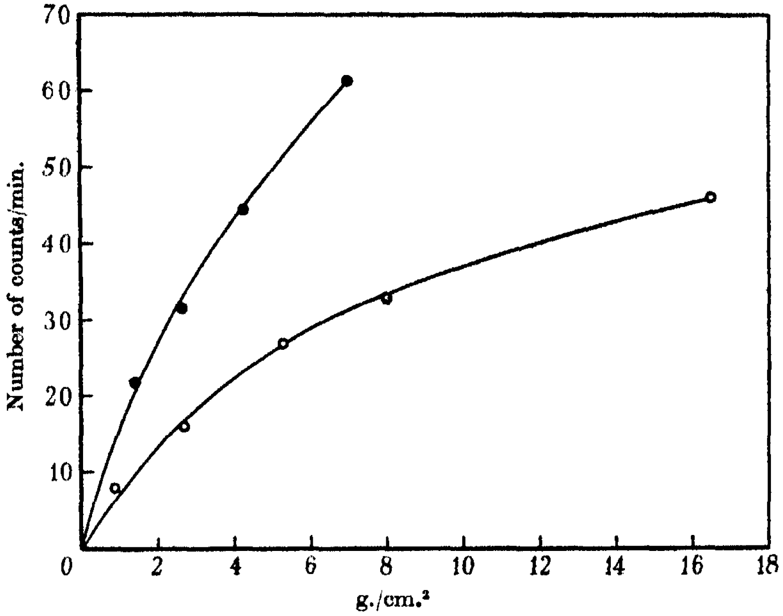


FIG. 6—Scattering curves for copper ○, and nickel ●.

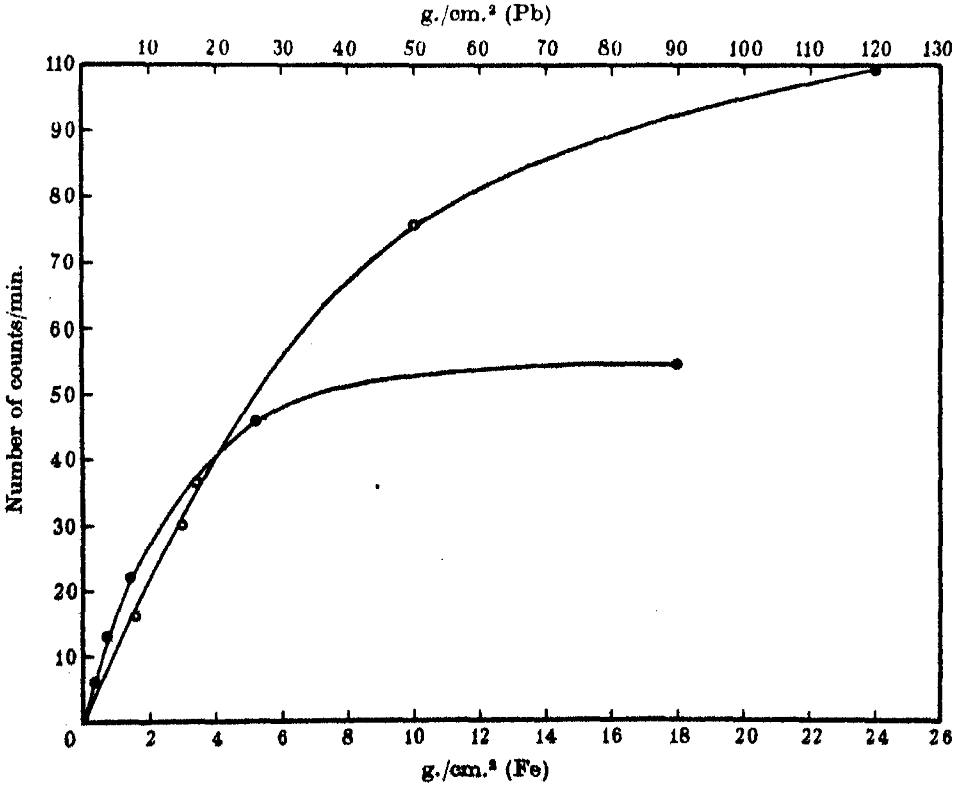


FIG. 7—Scattering curves for iron ○, and lead ●.

were obtained without the paraffin reflector round the BF_3 counter, the saturation value for thick paraffin therefore being only 200 instead of 300. For thick scatterers of carbon, nickel and lead, R_R was of the order of 5–10 counts/min., whereas we found $C_R \approx 0$ in these cases, from which we can conclude that R neutrons are not slowed down appreciably by exciting lattice vibrations in scatterers of comparatively high atomic weight.

The shape of the saturation curves is governed by the values of the scattering cross-section, the total cross-section, and by geometrical factors (densities of scatterers). It is in general not easy to interpret the whole saturation curve, but a definite meaning can be given to the initial slope (see § 4).

3—*Scattering of Slow Neutrons by Hydrogen bound in different Chemical Compounds*

We have compared the scattering of slow neutrons by water and paraffin both for thin layers, 1–2 mm. thick, and also for layers thick enough (about 3 cm.) to give the saturation values. For the thin layers three different sets of experiments were carried out, two comparing paraffin wax with water, and one comparing paraffin oil with water. In each comparison the thickness of water and paraffin were such as to contain equal numbers of hydrogen atoms, and the scatterers were adjusted so that their central planes were equidistant from A. Paraffin oil was regarded as consisting of "average" molecules $C_{18}H_{38}$. Alternate runs were made with the different scatterers. The results are given in Table II.

TABLE II

		Difference in %	Probable error in %
Thick layers	Paraffin wax—water	+1·5	2
Thin layers:			
Exp. No. 1	Paraffin wax—water	+2	2
„ No. 2	Paraffin wax—water	+0·6	1
„ No. 3	Paraffin oil—water	+3·0	2

From these results we can conclude that the scattering cross-section of hydrogen, whether bound in paraffin or in water molecules, does not vary by more than 2 %. A small positive difference between paraffin and water of about 1 % may be expected, since the scattering cross-section of carbon is about 15 % larger than that of oxygen (see below).

4—*Determination of Scattering Cross-Sections*

Our method of determining the scattering cross-sections of the elements is based on the assumption that the scattering of slow neutrons is, to a high approximation, spherically symmetrical. According to theory, this is true in a system in which the centre of gravity of the colliding particles is at rest, provided that the wave-length of the neutrons is large compared with nuclear dimensions (i.e. 10^{-13}–10^{-12} cm.), and that no resonance phenomena have to be considered. The wave-length of slow neutrons is of the order of 10^{-8} cm. For free nuclei, the scattering will be nearly spherically symmetrical in a system in which they are at rest if their mass is considerably larger than that of the neutron. Spherical symmetry will be more closely approximated, even for light nuclei, if they are chemically bound, and the neutron energies are small compared with the molecular vibration energies. Therefore, with the possible exception of the lightest element, hydrogen, which we shall discuss below, scattering would be practically spherically symmetrical in all cases. With our method of detecting slow neutrons resonance scattering would scarcely be noticeable because of the known small width of neutron resonance levels. For neutrons of room temperature effects due to interference should be negligible under our geometrical conditions.

Assuming, therefore, similar angular distribution of the scattered neutrons we can deduce the ratio of the scattering cross-sections of two elements from the initial slope of the saturation curves of the type shown in figs. 6 and 7. Let $I(w)$ be the number of counts per min. observed with a scatterer of w g./cm.^{2} of atomic weight a at A; then the ratio of the scattering cross-sections of the two elements will be given by

$$\frac{\sigma_1}{\sigma_2} = \frac{\left(\frac{aI_1}{dw_1}\right)_{\omega_1=0} a_1}{\left(\frac{dI_2}{dw_2}\right)_{\omega_2=0} a_2}.$$

As the initial part of the saturation curves is practically linear for small thicknesses, we can simplify the measurements of the relative cross-sections by measuring the number of counts per min., I°, for one thickness, w° g./cm.^{2}, of the scatterer, provided that I° is small compared with the saturation value. In order to correct roughly for the thickness ($2s$) in the case of thick scatterers which were used occasionally for elements with a small total cross-section we considered the whole scatterer concentrated at a plane distant s from the opening A. This was accomplished by introducing a correction function $F(s)$ defined by the ratio of the number of counts per min. obtained with a thin paraffin scatterer (1 mm. thick) at a distance s from A, divided by the number of counts per min. with the paraffin scatterer at A. For

$s = 0$, 3 and 6 mm. $F(s)$ was found to be 1, 0·91 and 0·83 respectively. Usually this geometrical correction was less than 10 %.

The final formula used for calculating the ratio of the scattering cross-sections of two elements from I^0, w^0 and s was

$$\frac{\sigma_1}{\sigma_2} = \frac{\dfrac{I_1^0}{w_2^0} a_1 F(s_1)}{\dfrac{I_1^0}{w_2^0} a_2 F(s_2)}.$$

As we have mentioned above carbon was chosen as the standard element. The scatterers investigated were compared directly with a graphite scatterer 5 mm. thick (0·79 g./cm.^{2}), and the cross-sections were calculated by assuming the scattering cross-section of carbon to be equal to the total cross-section determined previously, i.e. $\sigma_{\text{carbon}} = 4{\cdot}8_3 \times 10^{-24}$ cm.^{2}. For thin scatterers our neutron beam can be regarded effectively as consisting of C neutrons only, so that the cross-sections found correspond to the scattering of C neutrons.

In the case of hydrogen there is no need to determine the scattering cross-section by this method, as its cross-section for true absorption is known to be less than 1 % of its total cross-section. It is therefore sufficient to assume the scattering cross-section to be equal to the previously determined total cross-section. But it is interesting to compare the ratio of the initial slopes of the saturation curves of paraffin and carbon (which are reproduced in fig. 8 for a paraffin saturation value of 300/min.) with the ratio of the slopes of the absorption curves in the logarithmic plot of fig. 2. For this ratio we find 17·2 in the scattering case and 18·2 in the absorption case, the difference Δ, being about 6 %. One might expect a deviation in this sense as the scattering of C neutrons by protons bound in paraffin molecules is not quite spherically symmetrical, the forward direction being slightly favoured. According to Fermi (1936) the scattering probability per unit solid angle is proportional to $e^{-\frac{2E}{h\nu}(1-\cos\theta)}$, where θ is the angle at which the neutron has been scattered, E the neutron energy and $h\nu$ the oscillation energy of a hydrogen atom in the paraffin molecule. For C neutrons Fermi assumes $\frac{E}{h\nu} \approx \frac{1}{10}$, which with our geometrical conditions leads to a value $\Delta \approx 1$ %. The experimental value $\Delta \approx 6$ % could be explained by assuming $\frac{E}{h\nu} \approx \frac{3}{10}$.

Possible Errors—Apart from statistical errors which were usually of the order of 5–10 %, systematic errors in the determination of the scattering cross-sections may arise from hydrogen impurities, mainly in the form of water or occluded hydrogen, and from strongly absorbing impurities in the scatterers. The latter source of error was usually small, since comparatively pure chemicals were used in most cases. To reduce the effect of water impurities all suspected chemicals were carefully dried. Chemical compounds were avoided where feasible, since water impurities are less likely to occur with pure elements. Moreover, in the case of chemical compounds, the final accuracy is reduced by the introduction into the calculation of at least one further cross-section than that of carbon.

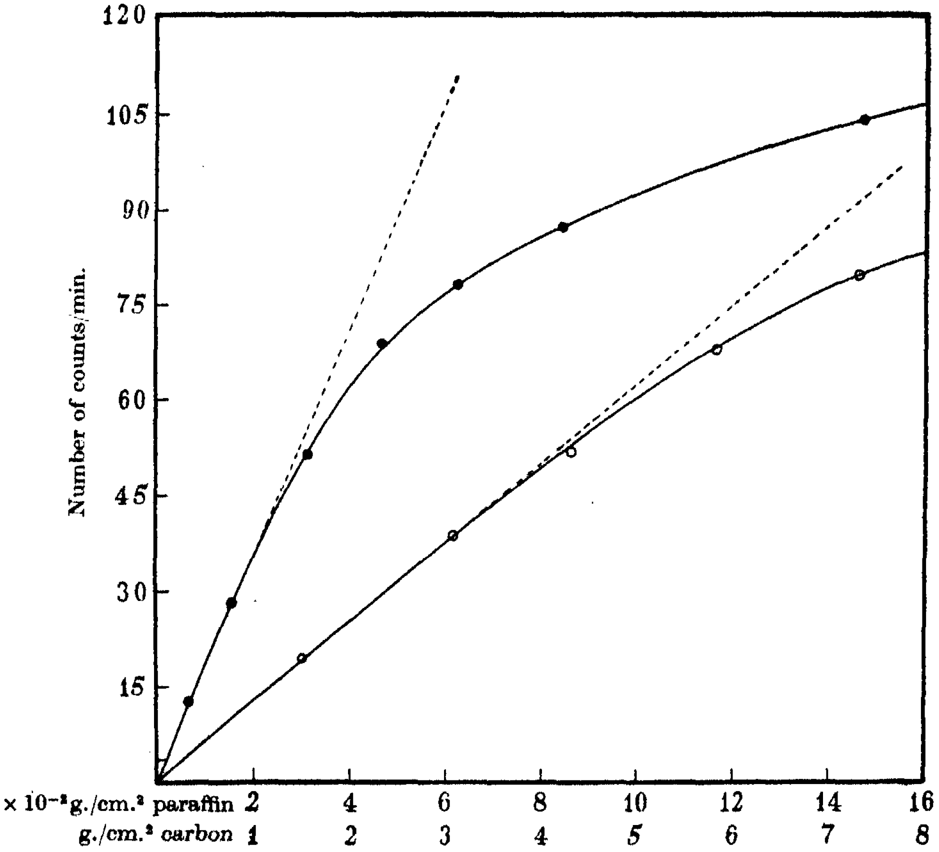


FIG. 8—Initial parts of the scattering curves for paraffin wax ○, and carbon ●.

RESULTS AND DISCUSSION

Our results are summarized in Table III.

In the cases where the number of counts observed with a scatterer, thick enough to absorb more than half of a beam of C neutrons, was less than

TABLE III

			$\sigma_{\text{total}} \times 10^{24}$ cm.$^{-2}$	
Element	Scatterer	$\sigma_{\text{scatt}} \times 10^{24}$ cm.$^{-2}$	Dunning and others	Present work (see part I)
1 H	C_nH_{2n+2}	*47·5*	35 (38)	47·5
1 D	D_2O	7	4·0	
2 He				
3 Li	Li	—	45	
4 Be	Be	6·9	5·3	
5 B	B	—	360	
6 C	C	*4·8*$_3$	4·1	4·8$_3$
7 N	NaN_3	8·2	11·3	
8 O	Al_2O_3	4·2	3·3	
9 F	CaF_2	4·1	2·5	
10 Ne				
11 Na	NaF	3·6	4·2	
12 Mg	Mg	4·2	3·5	
13 Al	Al	1·6	1·5	
14 Si	Si	1·7	2·5	
15 P	P	10·4	14·7	13·6
16 S	S	1·1	1·4	
17 Cl	KCl	< 10	39	
18 A				
19 K	KF	1·5	8·2	
20 Ca	Ca	9·5	11·0	
21 Sc				
22 Ti	TiO_2	6·2	11·9	
23 V	V	< 4	10	
24 Cr	Cr	3·6	4·9	
25 Mn	Mn	2·1	14·3	
26 Fe	Fe	10·3	12·0	13·6
27 Co	Co	~ 5	35	
28 Ni	Ni	12·4	15·4	19·7
29 Cu	Cu	8·6	7·5	11·9
30 Zn	Zn	5·2	4·7	
31 Ga	Ga	< 4		
32 Ge			Est. 75	
33 As	As	7·1	8·6	
34 Se	Se	12·7	19	20·4
35 Br	KBr	< 7	11·8	
36 Kr				
37 Rb				
38 Sr	SrS	9·5	Est. 9	
39 Y	Y_2O_3	< 14		
40 Zr	Zr	< 16	16·7	
41 Nb	Nb	5	Est. 14	
42 Mo	Mo	6·7	7·1	
43 Ma				
44 Ru	Ru	5·9	12·5	
45 Rh	Rh	—	115	
46 Pd	Pd	3	10	

TABLE III—(*continued*)

47 Ag	Ag	+	55	63
48 Cd	Cd	—	3300	
49 In	In	—		
50 Sn	Sn	4·9	4·0	
51 Sb	Sb	6·5	8·1	
52 Te	Te	5·3	8·2	
53 I	I	—	9·4	
54 X				
55 Cs				
56 Ba	BaF_2	8·2		
57 La			80	
58 Ce	Ce	(—)	25	
59 Pr			25	
60 Nd			220	
61 —				
62 Sm	Sm_2O_3	—	4700	
63 Eu			Est. 1000	
64 Gd	Gd_2O_3	—	30,000	
73 Ta	Ta	~10	27	
74 W	W	7·8	23	
75 Re	Re	—	89	
76 Os	Os	10·8	27	
77 Ir	Ir	+	285	
78 Pt	Pt	8·3	25	21·5
79 Au	Au	—	88	
80 Hg	Hg	+	380	
81 Tl	Tl	14·2	11	
82 Pb	Pb	12·9	8·6	12·5
83 Bi	Bi	8·9	8·2	8·9

2/min., this is indicated by the sign —. This does not necessarily imply that the absolute scattering cross-section is very small but it means that it is small (of the order of 1% or less) compared with the total cross-section of the element. In the cases where the number of counts was 2–5/min. for comparatively thick scatterers this is indicated by a + sign. Where the saturation values were 5–10/min. the accuracy obtained is rather low. This is indicated by the sign ~. With the exception of a few cases where we may not have escaped systematic errors due to impurities of hydrogen or strongly absorbing elements, the values for the scattering cross-sections should be reliable to 10–20%.

The values for the total cross-sections obtained by Dunning and others (1935) and the values quoted in part I, § 2 are given for comparison. The accuracy of the measurements is not yet sufficient to justify the calculation of the cross-sections for true absorption. The comparatively good agreement between the scattering cross-sections and the total cross-sections obtained in the cases of Pb and Bi which are known from various experiments to have a small cross-section for true absorption can be regarded as a verification of the assumption that the scattering is spherically symmetrical. When the scattering cross-section is the main part of the total cross-section our values for the scattering cross-sections are often larger than the values found by Dunning and others (1935) for the total cross-section. We have suggested a possible explanation for such a discrepancy in part I, § 2. The values given by Mitchell and others (1936) for the scattering cross-sections of some elements agree qualitatively with our values.

It can be seen from Table III that for the majority of the elements the cross-sections for scattering of slow neutrons lie between 1 and 15×10^{-24} cm.^{2}. A notable exception is hydrogen. However, according to Fermi (1936), the cross-section for scattering of C neutrons by free protons would be about 3·3 times smaller than that observed with chemically bound protons, i.e. about $14{\cdot}4 \times 10^{-24}$ cm.^{2}.

The cross-section for scattering of slow neutrons shows much less variation for different elements than does the cross-section for true absorption which varies by factors as high as 10^6 (between a few times 10^{-25} and a few times 10^{-19} cm.^{2}). The scattering cross-section of strong absorbers is $\leqslant / - 2\%$ of the total cross-section. This is in conformity with the theories of neutron scattering and absorption, due to Bohr, Breit and Wigner, and Bethe and Plazceck, in which these phenomena are treated as "many particle processes".

On the assumption that the interaction between a neutron and an electron can be represented by a positive or negative potential V of short range a, it can be shown (Condon 1936) that the contribution of all the electrons of an atom of charge Z to the scattering cross-section for slow neutrons will be approximately proportional to Z^2. We can therefore regard the minimum value of $\frac{\sigma_{\text{scatt}}}{Z^2}$ as an upper limit for the cross-section σ_{el} for scattering of slow neutrons by a bound electron. Assuming $\sigma_{\text{el}} < 10^{-24}$ cm.^{2}, Condon deduced for the absolute value of the constant $K = \frac{4\pi V a^3}{3}$ an upper limit $|K| < 30\,(mc^2)\,(e^2/mc^2)^3$. He points out that an electron-neutron interaction of this magnitude ($\frac{1}{6}$ to $\frac{2}{3}$ of this upper limit) could be the main cause of the isotopic displacement of spectral lines. From our measurements of σ_{scatt} we find the minimum value of $\frac{\sigma_{\text{scatt}}}{Z^2}$ to be of the order of 10^{-27} cm.^{2}, the smallest values being obtained for the elements palladium, iodine, gold and bismuth. With $\sigma_{\text{el}} < 10^{-27}$ cm.^{2} we obtain $|K| < mc^2(e^2/mc^2)^3$, which is too small to account for the observed isotopic displacement of spectral lines.

We wish to thank Lord Rutherford for the kind interest he has shown in our work. One of us (G. H. B.) wishes to express his gratitude to him for permission to work at the Cavendish Laboratory during leave of absence from the University of Sydney. We should also like to thank here all those who have kindly lent us chemicals used in this investigation, especially Captain Atkinson, of the Mond Nickel Company, who has kindly put at our disposal large quantities of the noble metals, Professor E. N. da C. Andrade, London, Professor R. S. Hutton, Cambridge, and Dr. J. K. Marsh, Oxford. One of us (M. G.) is indebted to the Department of Scientific and Industrial Research for a grant.

SUMMARY

Improved methods for measuring the "total" and the scattering cross-sections of the elements for slow neutrons are described. Determinations of the total cross-sections for a few elements and of the scattering cross-sections for the majority of the elements have been made and are summarized in tables. In contrast with the cross-section for true absorption which varies by factors of the order of 10^6 for different elements the scattering cross-section varies only by factors of the order of 10. The scattering of slow neutrons in hydrogen has been studied from various points of view (influence of chemical binding, slowing down of neutrons of medium velocity, etc.). It is found that elements heavier than hydrogen do not slow down neutrons of medium velocity appreciably. An upper limit for the cross-section for scattering of slow neutrons by electrons is found; it is about 10^{-27} cm.^{2}.

REFERENCES

Amaldi, D'Agostino, Fermi, Pontecorvo, Rasetti and Segrè 1935 *Proc. Roy. Soc.* A, **149**, 522–58.
Amaldi and Fermi 1936 *Phys. Rev.* **50**, 899–928.
Budnitzki and Kurtschatow 1935 *Phys. Z. Sowjet.* **8**, 170–8.
Condon 1936 *Phys. Rev.* **49**, 459–61.
Dunning, Pegram, Fink and Mitchell 1935 *Phys. Rev.* **48**, 265–80.
Fink 1936 *Phys. Rev.* **50**, 738–47.
Griffiths and Szilard 1937 *Nature, Lond.*, **139**, 323–4.
Hopwood and Chalmers 1935 *Nature, Lond.*, **135**, 341–2.
Mitchell and Murphy 1935 *Phys. Rev.* **48**, 653–6.
Mitchell, Murphy and Whitaker 1936 *Phys. Rev.* **50**, 133–7.
Pontecorvo and Wick 1936 *Ric. Sci.* **7** (1), 134–7 and 220–2.

Quaternion Treatment of the Relativistic Wave Equation

By A. W. Conway, F.R.S.

(*Received* 31 *May* 1937)

1—INTRODUCTION

A set of matrices can be found which is isomorphic with any linear associative algebra. For the case of quaternions this was first shown by Cayley (1858), but the first formal representation was made by Peirce (1875, 1881). These were two-row matrices, and the introduction of the four-row matrices of Dirac and Eddington necessitated the treatment of a wave function as a matrix of one row (as columns). Quaternions have been used by Lanczos (1929) to discuss a different form of wave equation, but here the Dirac form is discussed, the wave function being taken as a quaternion and the four-row matrices being linear functions of a quaternion. Certain advantages are claimed for quaternion methods. The absence of the distinction between outer and scalar products in the matrix notation necessitates special expedients (Eddington 1936). Every matrix is a very simple function of the fundamental Hamiltonian vectors α, β, γ, so that the result of combination is at once evident and depends only on the rules of combination of these vectors. At all stages the relationship of the different quantities to four-space is at once visible. The Dirac-Eddington matrices, the wave equation and its exact solution by Darwin, angular momentum operators, the general and Lorentz transformation, spinors and six-vectors, the current-density four-vector are treated in order to exhibit the working of this method.

S and V for scalar and vector products are used. Quaternions are denoted by Clarendon type, and all vectors are in Greek letters.

2—THE DIRAC-EDDINGTON MATRICES

The Hamiltonian unit vectors α, β, γ satisfy the equations

$$\alpha^2 = \beta^2 = \gamma^2 = -1, \quad \beta\gamma = -\gamma\beta = \alpha, \text{ etc.}$$

In terms of these any quaternion $\boldsymbol{\psi}$ can be expressed in the form

$$\psi_0 + \psi_x\alpha + \psi_y\beta + \psi_z\gamma.$$

Denoting by $f\psi$ a linear function of the quaternion ψ a special form of $f\psi$ is $\mathbf{p}\psi\mathbf{q}$, where p and q are quaternions. We get $f^2\psi = \mathbf{p}^2\psi\mathbf{q}^2$. So that if $\mathbf{p}^2$ and $\mathbf{q}^2$ are both equal to -1 we have a function f for which symbolically $f^2 = 1$. If i denote the scalar imaginary we can at once write down the fifteen Eddington-Dirac matrices in quaternion form

$$\left.\begin{matrix} i(\ )\alpha, & i(\ )\beta, & i(\ )\gamma, \\ \alpha(\ )i, & \beta(\ )i, & \gamma(\ )i, \\ \alpha(\ )\alpha, & \alpha(\ )\beta, & \alpha(\ )\gamma, \\ \beta(\ )\alpha, & \beta(\ )\beta, & \beta(\ )\gamma, \\ \gamma(\ )\alpha, & \gamma(\ )\beta, & \gamma(\ )\gamma. \end{matrix}\right\} \qquad (2{\cdot}1)$$

If $f'\psi$ denote $\mathbf{p}'\psi\mathbf{q}'$, we have

$$ff'\psi = \mathbf{p}\mathbf{p}'\psi\mathbf{q}'\mathbf{q}, \quad f'f\psi = \mathbf{p}'\mathbf{p}\psi\mathbf{q}\mathbf{q}'.$$

Hence if the pre-factors $\mathbf{p}$, $\mathbf{p}'$ commute and the post-factors $\mathbf{q}$, $\mathbf{q}'$ anti-commute (or vice versa), we have $ff' + f'f = 0$. With this principle we can at once pick out from (2·1) the different anti-commuting sets of five such as

$$i(\ )\alpha, \quad i(\ )\beta, \quad \alpha(\ )\gamma, \quad \beta(\ )\gamma, \quad \gamma(\ )\gamma.$$

It may be noticed that we can generalize (2·1) by replacing α, β, γ where they occur as pre-factors (or post-factors) by any other set of orthogonal unit vectors α', β', γ'.

3—THE RELATIVISTIC WAVE EQUATION

If $\nabla \equiv \alpha\frac{\partial}{\partial x} + \beta\frac{\partial}{\partial y} + \gamma\frac{\partial}{\partial z}$, the operators $\nabla(\ )\alpha$, $\frac{\partial}{c\partial t}(\ )\beta$, $mc(\ )\gamma$ anti-commute. Hence the operation $\nabla(\ )\alpha + \frac{\partial}{c\partial t}(\ )\beta + \frac{mc}{h}(\ )\gamma$ on being repeated gives

$$\left(-\nabla^2 - \frac{\partial^2}{c^2\partial t^2} - \frac{m^2c^2}{h^2}\right)(\ ).$$

(It is to be noted that $-\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$.)

So that we can take

$$\nabla\psi\alpha + \frac{\partial}{c\partial t}\psi\beta + \frac{mc}{h}\psi\gamma = 0 \qquad (3{\cdot}1)$$

as the wave equation for a free electron.

In a more expanded form this can be written

$$\left(\alpha\frac{\partial}{\partial x}+\beta\frac{\partial}{\partial y}+\gamma\frac{\partial}{\partial z}\right)\psi\alpha+\frac{\partial}{c\partial t}\psi\beta+\frac{mc}{h}\psi\gamma=0. \tag{3·2}$$

By using other linear operators we have different forms for this equation. For instance, Dirac's original form would be written, in quaternions,

$$\left(\alpha\frac{\partial}{\partial x}+\beta\frac{\partial}{\partial y}+i\frac{\gamma}{c}\frac{\partial}{\partial t}\right)\psi\alpha+i\frac{\partial}{\partial z}\psi\beta+mc\psi\gamma=0. \tag{3·3}$$

However, by a well-known theorem if ψ is a solution of (3·2), $\psi-i\gamma\psi\gamma$ is a solution of (3·3).

If we have a field of force given by a scalar potential P and a vector potential Π, we have the electric force

$$\epsilon=-\nabla P-\frac{1}{c}\frac{\partial\Pi}{\partial t} \tag{3·4}$$

and the magnetic force

$$\eta=\nabla\Pi-\frac{1}{c}\frac{\partial P}{\partial t}, \tag{3·5}$$

because $S\nabla\Pi-\frac{1}{c}\frac{\partial P}{\partial t}=0.$

We now write the equation

$$\left(\nabla+\frac{ie}{ch}\Pi\right)\psi\alpha+\left(\frac{\partial}{c\partial t}-\frac{ie}{ch}P\right)\psi\beta+\frac{mc}{h}\psi\gamma=0. \tag{3·6}$$

On repeating the operator we get by (3·4) and (3·5)

$$\left\{-\left(\nabla+\frac{ie}{ch}\Pi\right)^2-\left(\frac{\partial}{c\partial t}-\frac{ie}{ch}P\right)^2-\frac{m^2c^2}{h^2}-\frac{ie}{ch}\eta\right\}\psi-\frac{ie}{ch}\epsilon\psi\gamma=0.$$

Multiplying (3·1) by α from the right and making a slight change in the vectors we get

$$\nabla\psi+\frac{\partial}{c\partial t}\psi\alpha+\frac{mc}{h}\psi\beta=0. \tag{3·7}$$

We shall call this type of equation

$$\mathbf{H}\psi=0. \tag{3·7}$$

4—ANGULAR MOMENTUM OPERATORS

If ρ denote the position vector $\alpha x+\beta y+\gamma z$ we get

$$\nabla S\gamma\rho\nabla - S\gamma\rho\nabla.\nabla = V\gamma\nabla = \tfrac{1}{2}(\gamma\nabla-\nabla\gamma),$$

so that

$$(S\gamma\rho\nabla+\tfrac{1}{2}\gamma)\,\nabla = \nabla(S\gamma\rho\nabla+\tfrac{1}{2}\gamma).$$

Hence $S\gamma\rho\nabla+\frac{1}{2}\gamma = im'$, a scalar constant, is a Dirac integral of $\mathbf{H}\psi = 0$.

If ϕ is the azimuthal angle corresponding to the axis γ we thus have

$$\left(\frac{d}{d\phi}-\tfrac{1}{2}\gamma+im'\right)\psi = 0. \tag{4·1}$$

If ψ contains ϕ in the form $e^{is\phi}$

$$i\gamma\psi = -2(s+m')\,\psi.$$

The latent roots ("proper values") of the function $i\gamma(\ )$ are $+1$ and -1.

Hence

$$s+m' = \pm\tfrac{1}{2}.$$

To complete the solution if we have any linear function $f(\ )$ for which $f^2 = 1$, the axes are given by $f(\mathbf{q})+\mathbf{q}$, $f(\mathbf{q})-\mathbf{q}$, where $\mathbf{q}$ is *any* quaternion. Applying this to $i\gamma(\ )$ we get that

$$\psi = \{u(\alpha+i\beta)+v(\gamma-i)\}\,e^{-(m'+\frac{1}{2})i\phi}+\{u'(\alpha-i\beta)+v'(\gamma+i)\}\,e^{-(m'-\frac{1}{2})i\phi}, \tag{4·2}$$

where u, v, u', v' are scalars not involving ϕ.

We have

$$(V\rho\nabla+1)\rho = -\rho(V\rho\nabla+1),$$

$$(V\rho\nabla+1)\,\nabla = -\nabla(V\rho\nabla+1).$$

Hence

$$(V\rho\nabla+1)\rho\nabla = \rho\nabla(V\rho\nabla+1). \tag{4·3}$$

Also

$$(V\rho\nabla+1)\rho\psi\alpha\gamma = \rho(V\rho\nabla+1)\,\psi\gamma\alpha,$$

$$(V\rho\nabla+1)\rho\psi\beta\gamma = \rho(V\rho\nabla+1)\,\psi\gamma\beta.$$

Hence the operator $(V\rho\nabla+1)(\ )\gamma$ satisfies the relation

$$(V\rho\nabla+1)\rho\mathbf{H}\psi\gamma = \rho\mathbf{H}(V\rho\nabla+1)\,\psi\gamma. \tag{4·4}$$

We can also show that

$$(V\rho\nabla)^2+V\rho\nabla = -\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta}-\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}, \tag{4·5}$$

since $V\rho\nabla r = 0$, where $r^2 = x^2+y^2+z^2$.

The equations (4·1), (4·4) also hold when there is an electrostatic potential which is a function of r only.

5—The Hydrogen Atom

The wave equation has the form

$$\nabla\psi - \frac{i}{ch} W\psi\alpha + \frac{mc}{h}\psi\beta = 0, \quad W = E + \frac{e^2}{r}. \tag{5·1}$$

Let u be a solid harmonic function of order $l+1$,

therefore $$\nabla^2 u = 0 \quad \text{and} \quad \nabla^2(\nabla u r^{-(2l+1)}) = 0, \tag{5·2}$$

$$\nabla^2(\nabla u r^{-(2l+1)}) = -(2l+1)\,\nabla(\rho\nabla u r^{-(2l+3)}) = 0. \tag{5·3}$$

Hence if we substitute for ψ in (5·1) the expression

$$\psi = Q_1\nabla u(i\gamma+1) + Q_2\rho\nabla u r^{-(2l+3)}(i\alpha-\beta), \tag{5·4}$$

where Q_1, Q_2 are scalar functions of r, we get

$$\left\{\frac{\partial Q_1}{r\partial r} + r^{-(2l+3)}\left(-\frac{W}{ch}+\frac{mc}{h}\right)Q_2\right\}\rho\nabla u(i\gamma+1) \\ -\left\{\frac{\partial Q_2}{r\partial r} + r^{-(2l+1)}\left(\frac{W}{ch}+\frac{mc}{h}\right)Q_1\right\}\nabla u(i\alpha-\beta), \tag{5·5}$$

so that (5·4) is a solution of (5·1) if

$$\left.\begin{aligned} \frac{\partial Q_1}{r\partial r} + r^{-(2l+3)}\left(-\frac{W}{ch}+\frac{mc}{h}\right)Q_2 = 0, \\ \frac{\partial Q_2}{r\partial r} + r^{-(2l+1)}\left(\frac{W}{ch}+\frac{mc}{h}\right)Q_1 = 0. \end{aligned}\right\} \tag{5·6}$$

The substitutions $Q_1 = F_1 r^{-l}$, $Q_2 = F_2 r^{(l+2)}$ reduce these equations to the usual forms. It may be noticed that u is a general spherical harmonic of order n and we can reduce (5·4) to Darwin's solutions* by the substitution $u = r^{l+1}P^s_{l+1}(\mu)\,e^{is\phi}$ and making use of relations (Conway 1932–3) like the following:

$$\begin{aligned}\nabla r^{l+1} P^s_{l+1}(\mu)\,e^{is\phi} = &-\tfrac{1}{2}(\alpha-i\beta)\,r^l P^{s+1}_l(\mu)\,e^{i(s+1)\phi} \\ &+\tfrac{1}{2}(\alpha+i\beta)(l+s)(l+s+1)\,r^l P^{s-1}_l(\mu)\,e^{i(s-1)\phi} \\ &+\gamma(l+s+1)\,r^l P^s_l(\mu)\,e^{is\phi}.\end{aligned}$$

---

* Darwin (1928, p. 669). The Legendre functions used are slightly different.

6—COMPLEX CONJUGATE OF A QUATERNION

A quaternion $\mathbf{q}$ can always be written in the form $u+\sigma v$, where σ is a unit vector. The quaternion conjugate of $\mathbf{q}$ is

$$K\mathbf{q} = u - \sigma v,$$

and we have the following properties:

$$K(\mathbf{q}_1\mathbf{q}_2\cdots\mathbf{q}_n) = K\mathbf{q}_n K\mathbf{q}_{n-1}\cdots K\mathbf{q}_1, \tag{6·1}$$

$$\mathbf{q}K\mathbf{q} = u^2+v^2. \tag{6·2}$$

Hamilton called this latter quantity the norm of the quaternion, $N\mathbf{q}$, and its square root with the real part positive the tensor $T\mathbf{q}$.

The complex conjugate of a quaternion, $\mathbf{q}$, is the imaginary conjugate of its quaternion conjugate. Denoting it by $\mathbf{q}^*$ we have

$$\mathbf{q}^* = u^* - \sigma v^*, \tag{6·3}$$

where u^*, v^*, are the imaginary conjugates of the scalars u, v, the vector being regarded as real. Thus

$$\mathbf{q}\mathbf{q}^* = uu^* + vv^* + \sigma(vu^* - uv^*), \tag{6·4}$$

in which we may note that the scalar part is real and the coefficient of σ imaginary, and

$$(\mathbf{q}_1\mathbf{q}_2\cdots\mathbf{q}_n)^* = \mathbf{q}_n^*\mathbf{q}_{n-1}^*\cdots\mathbf{q}_1^*. \tag{6·5}$$

A real rotation of amount θ about an axis defined by the unit vector α is given by

$$\mathbf{q}' = \mathbf{r}\mathbf{q}\mathbf{r}^{-1}, \quad \mathbf{r} = \cos\tfrac{1}{2}\theta + \alpha\sin\tfrac{1}{2}\theta. \tag{6·6}$$

Here $\mathbf{r}^* = K\mathbf{r} = \mathbf{r}^{-1}, \quad N\mathbf{r} = 1.$

A Lorentz transformation corresponding to a velocity $c\tanh\theta$ along α is given by

$$\mathbf{q}' = \mathbf{l}\mathbf{q}\mathbf{l}, \quad \mathbf{l} = \cosh\tfrac{1}{2}\theta + i\alpha\sinh\tfrac{1}{2}\theta. \tag{6·7}$$

Here $\mathbf{l}^* = \mathbf{l}, \quad N\mathbf{l} = 1.$

The general orthogonal transformation in four-space is made up of any combination of the above two transformations and can be generally written

$$\mathbf{q}' = \mathbf{b}\mathbf{q}\mathbf{a}, \tag{6·8}$$

in which it is easy to show that

$$\mathbf{a}^* = \mathbf{b}, \quad \mathbf{b}^* = \mathbf{a}, \quad N\mathbf{a} = N\mathbf{b} = 1, \tag{6·9}$$

$$K\mathbf{a} = \mathbf{a}^{-1}, \quad K\mathbf{b} = \mathbf{b}^{-1} \quad \text{and} \quad K\mathbf{q}' = \mathbf{a}^{-1}K\mathbf{q}\mathbf{b}^{-1}. \tag{6·10}$$

7—GENERAL ORTHOGONAL TRANSFORMATION OF THE WAVE EQUATION

Writing $$\mathbf{q} = \alpha x + \beta y + \gamma z + ict, \quad \mathbf{D} = \nabla + \frac{i\partial}{c\partial t},$$

we have $$\mathbf{q}' = \mathbf{bqa}, \quad \mathbf{D}' = \mathbf{a}^{-1}\mathbf{D}\mathbf{b}^{-1}, \quad K\mathbf{D}' = \mathbf{b}K\mathbf{D}\mathbf{a},$$

and since $$2\nabla = \mathbf{D} - K\mathbf{D}, \ \frac{2i}{c}\frac{\partial}{\partial t} = \mathbf{D} + K\mathbf{D},$$

we get $$\left.\begin{aligned} 2\nabla &= \mathbf{a}\nabla'\mathbf{b} + \mathbf{b}^{-1}\nabla'\mathbf{a}^{-1} + i(\mathbf{ab} - \mathbf{b}^{-1}\mathbf{a}^{-1})\frac{\partial}{c\partial t'}, \\ 2\frac{\partial}{c\partial t} &= -i(\mathbf{a}\nabla'\mathbf{b} - \mathbf{b}^{-1}\nabla'\mathbf{a}^{-1}) + (\mathbf{ab} + \mathbf{b}^{-1}\mathbf{a}^{-1})\frac{\partial}{c\partial t'}. \end{aligned}\right\} \tag{7·1}$$

Substituting in the equation
$$\mathbf{H}\boldsymbol{\psi} \equiv \nabla\boldsymbol{\psi} + \frac{\partial}{c\partial t}\boldsymbol{\psi}\alpha + \frac{mc}{h}\boldsymbol{\psi}\beta,$$

we get $$\mathbf{H}\boldsymbol{\psi} = \tfrac{1}{2}(\mathbf{a} + \mathbf{b}^{-1})\mathbf{H}'\boldsymbol{\psi}' - \tfrac{1}{2}(\mathbf{a} - \mathbf{b}^{-1})i\mathbf{H}'\boldsymbol{\psi}'\alpha, \tag{7·2}$$

where $$\mathbf{H}'\boldsymbol{\psi}' = \nabla'\boldsymbol{\psi}' + \frac{\partial}{c\partial t'}\boldsymbol{\psi}'\alpha + \frac{mc}{h}\boldsymbol{\psi}'\beta, \tag{7·3}$$

$$\boldsymbol{\psi}' = \tfrac{1}{2}(\mathbf{b} + \mathbf{a}^{-1})\boldsymbol{\psi} - \tfrac{1}{2}i(\mathbf{b} - \mathbf{a}^{-1})\boldsymbol{\psi}\alpha, \tag{7·4}$$

or $$\boldsymbol{\psi} = \tfrac{1}{2}(\mathbf{a} + \mathbf{b}^{-1})\boldsymbol{\psi}' + \tfrac{1}{2}i(\mathbf{a} - \mathbf{b}^{-1})\boldsymbol{\psi}'\alpha. \tag{7·5}$$

(7·4) and (7·5) give the general orthogonal transformation of wave functions.

If there is a field of electric force given by the potentials Π, P the equation $S\nabla\Pi - \frac{1}{c}\frac{\partial P}{\partial t} = 0$ may be written
$$S\left(\nabla + \frac{i\partial}{c\partial t}\right)(\Pi + iP) = 0.$$

Hence $\Pi + iP$ is a contravariant for the general transformation and $\Pi - iP$ is cogredient with $\nabla + \frac{i\partial}{c\partial t}$ and hence the above results (7·4) and (7·5) hold in this case also.

8—Spinors and Six-vectors

From (7·4) and (7·5) we get

$$\boldsymbol{\psi}' - i\boldsymbol{\psi}'\alpha = \mathbf{b}(\boldsymbol{\psi} - i\boldsymbol{\psi}\alpha), \tag{8·1}$$

$$\boldsymbol{\psi}' + i\boldsymbol{\psi}'\alpha = \mathbf{a}^{-1}(\boldsymbol{\psi} + i\boldsymbol{\psi}\alpha). \tag{8·2}$$

For the Lorentz transformation $\mathbf{b} = \mathbf{a} = \mathbf{1}$. Hence $\boldsymbol{\psi} - i\boldsymbol{\psi}\alpha$ is a spinor of the first class. It may be noted that if

$$\boldsymbol{\psi} = \psi_x\alpha + \psi_y\beta + \psi_z\gamma + \psi_0,$$

$$\begin{aligned}\boldsymbol{\psi} - i\boldsymbol{\psi}\alpha &= (\psi_x - i\psi_0)\,\alpha + (\psi_y - i\psi_z)\,\beta + (\psi_z + i\psi_y)\,\gamma + \psi_0 + i\psi_x \\ &= (\psi_x - i\psi_0)\,(\alpha + i) + (\psi_y - i\psi_z)\,(\beta + i\gamma).\end{aligned}$$

Let ϕ and ψ be any two spinors so that

$$\boldsymbol{\psi}'(1 - i\alpha) = \mathbf{l}\boldsymbol{\psi}(1 - i\alpha),$$

$$\boldsymbol{\phi}'(1 - i\alpha) = \mathbf{l}\boldsymbol{\phi}(1 - i\alpha).$$

Taking the complex conjugate of the second equation we get

$$(1 - i\alpha)\,\boldsymbol{\phi}'^{*} = (1 - i\alpha)\,\boldsymbol{\phi}^{*}\mathbf{l}.$$

Multiplying and noting that $(1 - i\alpha)^2 = 2(1 - i\alpha)$ we get

$$\boldsymbol{\psi}'(1 - i\alpha)\,\boldsymbol{\phi}' = \mathbf{l}\boldsymbol{\psi}(1 - i\alpha)\,\boldsymbol{\phi}\mathbf{l},$$

so that from any two spinors we get a vector (Whittaker 1937) which obeys the Lorentz group

$$\boldsymbol{\psi}(1 - i\alpha)\,\boldsymbol{\phi}.$$

In the same way for the general transformation (8·1) and (8·2) we can prove that

$$\boldsymbol{\psi}'(1 - i\alpha)\,\boldsymbol{\psi}'^{*} = \mathbf{b}\{\boldsymbol{\psi}(1 - i\alpha)\,\boldsymbol{\psi}^{*}\}\,\mathbf{a},$$

and that $$\mathbf{D}'\boldsymbol{\psi}'(1 - i\alpha)\,\boldsymbol{\psi}^{*\prime} = \mathbf{a}^{-1}\{\mathbf{D}\boldsymbol{\psi}(1 - i\alpha)\,\boldsymbol{\psi}^{*}\}\,\mathbf{a},$$

so that $$S\mathbf{D}\boldsymbol{\psi}(1 - i\alpha)\,\boldsymbol{\psi}^{*} \text{ is an invariant.} \tag{8·3}$$

If $\mathbf{P} = \Pi + iP$ then $$\mathbf{DP} = \left(\nabla + \frac{i\partial}{c\partial t}\right)(\Pi + iP)$$

$$= \eta - i\epsilon = i(\epsilon + i\eta) \text{ the electromagnetic six-vector.}$$

For the general transformation (Conway 1911; Silberstein 1912, 1913)

$$\mathbf{D'P'} = \mathbf{b}^{-1}\mathbf{D}\mathbf{a}^{-1}\mathbf{a}\mathbf{P}\mathbf{b} = \mathbf{b}^{-1}\mathbf{D}\mathbf{P}\mathbf{b},$$

therefore

$$\epsilon' + i\eta' = \mathbf{b}^{-1}(\epsilon + i\eta)\,\mathbf{b},$$

which is a rotation, in general, complex.

As another example Dirac's six-component angular momentum in our notation is $V\mathbf{qD}$, and for the transformation we find

$$V\mathbf{q'D'} = V\mathbf{aqbb}^{-1}\mathbf{D}\mathbf{a}^{-1} = \mathbf{a}V\mathbf{qDa}^{-1}.$$

There are thus two different types of six-vectors arising from products of two quaternions.

9—Electric and Current Density

The fundamental wave equation may be written

$$\nabla\boldsymbol{\psi} + \frac{\partial}{c\partial t}\boldsymbol{\psi}\alpha + \frac{mc}{h}\boldsymbol{\psi}\beta = \frac{ie}{ch}(P\boldsymbol{\psi}\alpha - \Pi\boldsymbol{\psi}). \tag{9·1}$$

Its conjugate is

$$-\boldsymbol{\psi}^*\nabla - \alpha\frac{\partial}{\partial t}\boldsymbol{\psi}^* - \frac{mc}{h}\beta\boldsymbol{\psi}^* = -\frac{ie}{h}(-P\alpha\boldsymbol{\psi}^* + \boldsymbol{\psi}^*\Pi). \tag{9·2}$$

Operating on (9·1) with $\boldsymbol{\psi}^*(\ \ )\alpha$ and multiplying (9·2) into $\psi\alpha$ we get on subtraction and taking the scalar of the result

$$S\boldsymbol{\psi}^*\nabla\boldsymbol{\psi}\alpha - \frac{\partial}{c\partial t}S\boldsymbol{\psi}\boldsymbol{\psi}^* = 0,$$

or

$$-\frac{\partial}{c\partial t}S\boldsymbol{\psi}\boldsymbol{\psi}^* + S(\nabla V\boldsymbol{\psi}\alpha\boldsymbol{\psi}^*) = 0. \tag{9·3}$$

This is the continuity equation for the quaternion or the equivalent four-vector $-S\boldsymbol{\psi}\boldsymbol{\psi}^* + iV\boldsymbol{\psi}\alpha\boldsymbol{\psi}^*$, which, when normalized, is Darwin's four-vector of electric charge and current (Darwin 1928).

To obtain the usual form of this four-vector, if we put

$$\boldsymbol{\psi} = i(\psi_1 + \psi_3)\,\alpha + (\psi_3 - \psi_1)\,\beta - (\psi_2 + \psi_4)\,\gamma + i(\psi_2 - \psi_4)$$

we get

$$-S\boldsymbol{\psi}\boldsymbol{\psi}^* = 2(\psi_1\psi_1^* + \psi_2\psi_2^* + \psi_3\psi_3^* + \psi_4\psi_4^*),$$

$$V\boldsymbol{\psi}\alpha\boldsymbol{\psi}^* = 2\alpha(\psi_1\psi_3^* + \psi_1^*\psi_3 - \psi_2\psi_4^* - \psi_2^*\psi_4) + \text{etc.}$$

(9·3) can be written

$$S\left(\nabla + \frac{i\partial}{c\partial t}\right)(-S\boldsymbol{\psi}\boldsymbol{\psi}^* + iV\boldsymbol{\psi}\alpha\boldsymbol{\psi}^*) = 0. \tag{9·4}$$

Hence $-S\psi\psi^* + iV\psi\alpha\psi^*$ is a contravariant four-vector for the general orthogonal transformation.

In the last section we had the invariant

$$S\mathbf{D}\psi(1-i\alpha)\psi^* = S\left(\nabla + \frac{i\partial}{c\partial t}\right)\{\psi(1-i\alpha)\psi^*\}$$

$$= S\left(\nabla + \frac{i\partial}{c\partial t}\right)(S\psi\psi^* - iV\psi\alpha\psi^*)$$

$$+ S\left(\nabla + \frac{i\partial}{c\partial t}\right)(V\psi\psi^* - iS\psi\alpha\psi^*).$$

Hence we have another contravariant four-vector

$$V\psi\psi^* - iS\psi\alpha\psi^*.$$

REFERENCES

Cayley, A. 1858 *Philos. Trans.* **148**, 32.
Conway, A. W. 1911 *Proc. R. Irish Acad.* A, **29**, 1.
— 1932–3 *Proc. R. Irish Acad.* A, **41**, 12.
Darwin, C. G. 1928 *Proc. Roy. Soc.* A, **118**, 654.
Eddington, A. 1936 "Relativity Theory of Protons and Electrons." Camb. Univ. Press, 16.
Lanczos, von C. 1929 *Z. Phys.* **57**, 447–73.
Peirce, B. 1875 *Proc. Amer. Acad. Arts Sci.* **2**, 10.
— 1881 *Amer. J. Math.* **4**, 216–221.
Silberstein, L. 1912 *Phil. Mag.* **23**, 790.
— 1913 *Phil. Mag.* **25**, 135.
Whittaker, E. T. 1937 *Proc. Roy. Soc.* A, **158**, 41.

Theory of Scattering of Protons by Protons

BY SIR ARTHUR EDDINGTON, F.R.S.

(*Received* 24 *June* 1937)

1—In a series of papers in these *Proceedings* between 1928 and 1935, I have developed a combination of relativity theory and wave mechanics applicable to a number of fundamental problems. A more mature treatment is given in my book "Relativity Theory of Protons and Electrons" (1936) hereinafter referred to as *P. and E.* This theory will here be applied to the problem of close encounters of two protons. In particular it will be used to calculate the form and depth of the potential well which is found to occur in the wave equation of the relative motion.

As the problem is of a somewhat different type from those treated in my book, the main difficulty has been to formulate it in terms of the relativistic conceptions there introduced. After that has been accomplished, the mathematical calculations are comparatively elementary.

The calculation shows (§ 13) that there will be an energy term in the relative wave equation of amount

$$-Ae^{-r^2/k_0^2}, \tag{1·1}$$

where

$$A = 39{\cdot}2m_e c^2, \quad k_0 = 2{\cdot}20 . 10^{-13}\,\text{cm}. \tag{1·2}$$

A potential well with the same value of k_0 occurs also in the close encounter of two electrons; but its strength A is smaller in the ratio m_e/m_p, so that it is much less important in comparison with the Coulomb potential. We find no corresponding potential well in the encounter of an electron and a proton.

It is generally believed that the potential well in this problem is fundamentally the same as that which occurs in nuclear theory. The investigation is therefore of special importance as a connecting link between extra-nuclear and nuclear theory. This aspect is briefly discussed in § 14.

2—In problems treated by quantum theory a few particles are mentioned specifically, but the particles constituting "the rest of the universe" must be understood to be present also; for it would be idle to make predictions which could only be verified experimentally by first annihilating nearly all the matter that exists. It makes no essential difference if we try to limit the system under consideration by enclosing it within a boundary; that would merely give us the additional trouble of working out boundary conditions which shall be mathematically equivalent to the suppressed matter beyond the boundary. The current procedure in quantum theory apparently ignores the particles forming the rest of the universe; and an important part of my relativistic theory is concerned with the precise conditions which must be postulated in order that this procedure may be valid (*P. and E.* especially Chapters XI and XIV).

The main consideration is that energy and momentum can be represented alternatively by wave functions (as in quantum theory) or by curvature of space-time (as in relativity theory). It is therefore permissible to represent the energy and momentum of a few specifically mentioned particles by wave functions which occupy a space-time whose curvature represents the momentum and energy of the remaining particles.

We shall call the system whose energy is represented by wave functions the *microscopic system*, and the system whose energy is represented by curvature the *molar system*. Ideally the division can be made in any way we like; and I have found (*P. and E.* Chapter XIV) that the cosmical constant can be calculated by treating the entire universe firstly as a wholly molar system and secondly as a wholly microscopic system. But in more specialized problems it is expedient to make the separation in such a way that the formulae of current quantum theory can be used as far as possible. The simplest "field-free" formulae presuppose a definite form of molar system.

When, as usual, we assume that our wave functions occupy a uniform static space, we implicitly postulate that the molar system has the energy tensor which corresponds to uniform static space. In other words the unspecified particles form an Einstein universe. Expressing the same condition from the point of view of quantum theory, they form a system in its ground state with the energy levels fully occupied up to a certain limit energy (*P. and E.* § 14·3). The individually specified particles forming the microscopic system must then occupy vacant levels above the limit level. In these idealized conditions it is possible to treat the microscopic system without explicit reference to the rest of the universe, and the procedure of current quantum theory is valid; the only effect of the unspecified particles is that, by filling the energy levels up to the limit level, they fix a threshold energy for the particles of the microscopic system. Reverting to the language of relativity theory, they provide a pure inertial field for the specified particles; this inertial field is represented by the $g_{\mu\nu}$ of the space whose curvature is defined by their energy tensor.

We are not putting forward a hypothesis as to the actual state or structure of the universe. We are formulating the conditions in which the simple field-free equations of quantum theory would be precisely fulfilled. Deviations from the ideal state of uniformity and non-excitation of the background are provided for by introducing gravitational, electromagnetic or radiation fields into the microscopic system.

Wide departures from the ideal conditions may be found to have little influence on the observable results. For example, the actual universe is found to be far from the Einstein equilibrium state; but the gravitational-cum-cosmical field representing the deviation has negligible effect, except on systems on a cosmical scale. But it does not follow that these departures have negligible effect on the intermediate formulae. Some of the most frequently used formulae refer to distributions on a cosmical scale, e.g. "infinite" plane waves. It would make negligible difference to the scattering of two protons if they were assumed to be in a de Sitter world instead of an Einstein world; but if we desired to treat the problem in a de Sitter world, it would be necessary to use formulae adapted to a de Sitter world from the beginning. The kind of mistake we have to guard against is integrating (or normalizing) over a de Sitter world the current formulae which presuppose an Einstein world.

Primarily we include "the rest of the universe" because it is impossible to annihilate it; and formulae which did not take account of it, explicitly or implicitly, would have no application to practical problems. But it also appears in the theory in another aspect, namely as a substitute for the reference objects with respect to which the positions and momenta of the individually specified particles would be measured in actual experiments. Unless such a substitute is provided the formulae must again fail to have any observable application; for the measured quantities, to which they refer, would be non-existent (*P. and E.* § 11·1).

3—For our scattering problem we have to determine a set of statistically steady distributions (eigenstates), nominally of two protons, but actually of all the particles of the universe including two specified protons. In accordance with § 2, the whole steady system is treated as a superposition of two independently steady systems, namely a microscopic and a molar system.

Even in the simpler problem of an electron and proton (instead of two protons) an adjustment of the frontier is necessary before separating the two systems. If two particles were withdrawn from the Einstein universe it would no longer have the mass or energy necessary for self-equilibrium. Thus we must continue to count the original energy of the proton and electron as energy of the molar system to be represented by space-time curvature. The energy represented by wave functions is therefore not the whole energy of the proton and electron, but the excess over that which they would have had as unspecified neutral particles at the limit level. This adjustment, which leads to the quadratic equation determining the masses of the proton and electron (*P. and E.* §§ 11·6, 12·6), has been absorbed into the practical definition of rest mass, and is thereby automatically provided for.

The *liaison* between my relativistic theory and current quantum theory has been provided by a detailed investigation of the hydrogen atom; and accordingly the theory in *P. and E.* has been developed primarily for application to systems which, like the hydrogen atom, contain equal numbers of positive and negative charges. When the object system contains an excess of charges of one sign, a new point arises which has not hitherto received attention. In selecting two protons to form an object system we leave the rest of the universe negatively charged; and there will be a mutual negative electrical energy of the positively charged object system and the negatively charged rest of the universe. When the whole is in a steady state, neither part is in self-equilibrium, but is only held in equilibrium by interaction with the other part. This violates the understanding that the entire universe is to be separated into two independently steady systems. Without some adjustment of the energy it would not be legitimate to apply the formulae of current quantum theory.

For a microscopic system consisting of a proton and electron, there is no electric interaction between the neutral system and the rest of the universe which is also neutral. When this is changed to a system of two protons, we have to account for a "redundant" energy Ω in the microscopic system, equal to *twice* the mutual energy of the two protons; since it is the replacement of the electron by a second proton which disturbs the balance. Each proton may therefore be considered to have a redundant energy $\frac{1}{2}\Omega$. The sense in which Ω is redundant will be examined in § 9.

Theoretically the same point arises when the microscopic system is a single proton, since the rest of the universe is left negatively charged. But formulae for a single proton are for the most part only of abstract interest; and it is scarcely necessary to consider an electric charge which is given no opportunity to manifest itself.

It may be suggested that it would be sufficient to treat the two protons as moving in a uniform field of negative potential arising from the negative charge of the rest of the universe. If this were so, there would be no effects of any interest. But the usual field treatment has limited scope, since it represents the action of the molar system on the microscopic system and not the action of the microscopic system on the molar system. We cannot here neglect the latter interaction. It should be remembered that the usual one-way treatment of interaction is a convention rather than a practical approximation. A proton normally polarizes surrounding matter; and the consequent redistribution of charge creates a well of negative potential in which the proton lies. The resulting negative energy (Debye-Hückel energy) is by no means unimportant practically; but there is no place for it in the wave equation of the proton, which provides only for rigid electromagnetic fields. In our scattering problem we are trying to discover the origin of a well of negative potential, which is known to occur observationally. We are not likely to find it unless we use a treatment more rigorous than the rigid field method.

Before proceeding further we calculate the redundant energy Ω.

4—We begin by calculating the average potential throughout spherical space due to a point charge ϵ. A form of metric of spherical space is

$$ds^2 = -(1-r^2/R^2)^{-1}dr^2 - r^2d\theta^2 - r^2\sin^2\theta\, d\phi^2 + dt^2. \tag{4·1}$$

The radial force is then (Eddington 1923, equation (78·4))

$$-\frac{\partial\kappa}{\partial r} = \frac{\epsilon}{r^2\sqrt{(1-r^2/R^2)}}, \tag{4·2}$$

whence, by integration,

$$\kappa = \epsilon\sqrt{(1-r^2/R^2)}/r. \tag{4·3}$$

Multiplying by the volume element

$$dV = \frac{r^2\sin\theta\, d\theta\, d\phi\, dr}{\sqrt{(1-r^2/R^2)}}$$

we have

$$\kappa dV = \epsilon r \sin\theta\, d\theta\, d\phi\, dr. \tag{4·4}$$

Integrating over a hemisphere,* the limits of r are 0 and R; we obtain

$$\int \kappa dV = 2\pi R^2\epsilon. \tag{4·5}$$

If the solution (4·3) is continued into the other hemisphere, we find that there is also a charge ϵ at the antipodal point. If therefore we double (4·5) to give the integral over the whole sphere, we must halve it again to obtain the part attributable to one charge ϵ. Hence if V is the whole volume, and $\bar{\kappa}$ the average potential throughout all space due to one point-charge ϵ,

---

* In the complete hypersphere r goes from 0 to R and back to 0. We halve it by taking half this range. Thus the remaining co-ordinates θ, ϕ, have their full ranges 0 to π, and 0 to 2π.

$$V\bar{\kappa} = 2\pi R^2\epsilon. \tag{4·6}$$

Since $V = 2\pi^2R^3$,
$$\bar{\kappa} = \epsilon/\pi R. \tag{4·7}$$

If another charge ϵ has an even probability distribution throughout space, the mutual energy is

$$\epsilon\bar{\kappa} = \epsilon^2/\pi R. \tag{4·8}$$

This applies also if one charge has an arbitrary probability distribution and the other an even probability distribution, if the two distributions are uncorrelated. As explained in § 3, the redundant energy is twice this mutual energy, namely

$$\Omega = 2\epsilon^2/\pi R. \tag{4·9}$$

5—As in the classical problem of two bodies, the motion of the two protons is analysed into an external motion (motion of the centre of mass) and an internal motion (relative orbit). Correspondingly, their combined wave function is analysed into external and internal wave functions. In the usual semi-classical terminology the internal wave function is said to represent a particle (called the internal particle) of reduced mass $\frac{1}{2}m_p$ whose co-ordinates ξ_μ are the relative co-ordinates $(x_2)_\mu - (x_1)_\mu$ of the two protons in a space-time frame whose time direction agrees with the momentum vector of the external particle.

It may be noted that the most essential difference between the problem of two protons and the problem of a proton and electron is that in the former the external particle has charge 2ϵ and in the latter it is neutral.

Since in the frame to which the internal wave functions are referred the external particle has an exact momentum (zero), it has entirely uncertain position; that is to say, the centre of mass has uniform probability distribution throughout space. The specialized relative distribution, prescribed by the internal wave function, is smeared evenly over space by the external wave function. It may save misconception if we bear in mind that the wave functions introduced in the ordinary treatment of the scattering problem (which are internal wave functions) specify *correlations* of the probability distributions of the two protons; but the protons individually continue to have even probability distribution throughout space—just like the unspecified protons and electrons.

The question arises, How can a system with vector characteristics be uniformly distributed over spherical space? Density and other scalar characteristics can be uniformly distributed; but there is no such thing as a uniform vector field in curved space. An encounter of two protons is associated with a particular direction of relative motion (indicated in the internal wave function); any attempt to transfer the encounter to another point in spherical space is frustrated by our inability to lay down a parallel or corresponding direction for the new wave function.

This is an example of the difficulties which are created by the replacement in current quantum theory of the double wave functions (which contain the observables) by abstract *simple* wave functions—a procedure which can be justified in some applications, but continually breaks down when we extend the range of problems. To avoid the difficulty we must revert to the double wave functions. These will represent the microscopic distribution fixed in a geometrical frame, coupled with an even probability distribution of orientation of the molar system which brings each point of it to the geometrical origin in turn. In this way we reconcile the two conditions that the distribution with vector characteristics is undisplaced in the geometrical frame, although relative to the other particles of the universe it has uniform probability distribution. The molar system admits of uniform distribution because it is purely scalar, vector characteristics having disappeared by averaging.

Let Ψ be the double vector wave function representing the combined distribution of an elementary object particle and a similar comparison particle. Although actual factorization is generally impossible, it is convenient to write it as a *symbolic* product $\Psi = \psi(x)\,\phi(x')$, where ψ and ϕ refer to the object particle and comparison particle, respectively. We then substitute for the single comparison particle an average of the N particles constituting the molar system, so as to obtain the combined wave function of the object particle and the molar system. It is assumed that ϕ reduces to a scalar, constant throughout space. We can then separate off the constant (along with its share of the periodic time factor), leaving $\psi(x)$ which exhibits the object particle as though it were detached from the rest of the universe and existed only in a geometrical frame. The energy corresponding to the part of the periodic time factor removed along with ϕ is the fundamental mass unit m_0 (*P. and E.* equations (12·61), (14·71)).

When this is extended to an object system consisting of two protons, four double wave functions are introduced, namely,

$$\Psi_1 = \psi_1(x_1)\,\phi_1(x_1'), \quad \Psi_2 = \psi_2(x_2)\,\phi_2(x_2'), \tag{5·1}$$

$$\Psi_i = \psi_i(\xi)\,\phi_i(\xi'), \quad \Psi_e = \psi_e(\mathrm{x})\,\phi_e(\mathrm{x}'), \tag{5·2}$$

referring respectively to the first proton, second proton, internal particle, external particle, with comparison particles of similar character. The single comparison particles are subsequently replaced by molar averages. We have

$$\xi = x_2 - x_1, \quad \xi' = x_2' - x_1', \quad \mathrm{x} = \tfrac{1}{2}(x_1 + x_2), \quad \mathrm{x}' = \tfrac{1}{2}(x_1' + x_2'). \qquad (5\cdot3)$$

In saying above that ϕ reduces to a constant scalar, we are not making a hypothesis but formulating the condition that the current procedure (which assumes that $\psi(x)$ can be separated) is valid. We must therefore examine the process of averaging closely in order to see how far the separation is justified; more especially the case of ϕ_i must be examined, since it is not quite comparable with the others.

6—Considering first a single object particle, the principal points to be noticed are as follows:

(1) The first step is to replace the N particles with charge and spin by neutral particles. I think it would in any case be illegitimate to assume independent probability distributions of positive and negative particles, neglecting the electric interaction which prevents any appreciable local inequality of their distribution. (The strongest electric charges encountered in practice constitute an inequality which is insignificant from the statistical point of view.) But the primary reason for considering neutral particles is that the molar system is the idealized substitute for the material objects which would be used for reference purposes in actual experiments; if these were charged, corrections equivalent to removing the charge would be applied to the measurements. The effect of retaining charged particles would be to give us a reference system, not only unsatisfactory to the experimenter, but unlike any material system in our experience, and unrepresentable in the Riemannian space of our experience.

(2) We next reduce the N neutral particles to one equivalent particle. Here the treatment bifurcates according as we are dealing with momenta or positions. In *P. and E.* we were concerned with momenta. In that case the wave function ϕ of the rest of the universe, which is initially an N-tuple wave function, is transformed by second quantization into symbolically additive wave functions (*P. and E.* §§ 16·1–16·4). Introducing the condition (already found) that the molar system is in its ground state, the symbolic addition reduces to ordinary addition; and ϕ is reduced to a set of standing waves whose energy tensors are additive. The total energy can be calculated, and the part of the periodic time factor associated with ϕ is thereby determined. In this way the molar system is reduced to an equivalent particle of energy m_0. The expectation value of its momentum (in three dimensions)

is zero; but the mean square value of the momentum is such as to correspond to the energy m_0. No co-ordinates are assigned to the equivalent particle, since the investigation does not concern itself with position.

(3) We note that, since the expectation value of the momentum is zero, m_0 measures an *uncertainty* of momentum. Also since current quantum theory takes cognizance only of the simple wave functions, and treats them as primitive data, it presupposes that the reduction from double to simple wave functions has been made independently for each object particle, before any consideration of their relative probability distribution arises. Physically this means that when we measure the momentum of a particle relative to the material reference system, we do not define the particular frame within the blur of uncertainty which is employed; and if the momentum of a second particle is measured, it is referred to an independent frame within the blur. The uncertainty of the frame is therefore *not* eliminated when we take the difference between the measured momenta or positions of two particles.

(4) A different method of combination must be used in treating positions, since the second quantization in (2) eliminates those characteristics of the molar system which would be used as reference marks for position. When an ordinary reference body is employed, its centroid (or some other invariantly defined combination of the co-ordinates of its particles) provides the material origin from which the relative position of the object particle is measured. A measurement of the relative position creates a wave packet in the double wave function $\Psi = \psi(x)\,\phi(x')$ representing the combined probability distribution of the object particle and the centroid; so that all observational information relates directly to the double wave function. As before we have to make a special theoretical calculation of $\phi(x')$ in order to detach it from $\psi(x)$. In dealing with momenta it was the periodic time factor in ϕ which was required; in dealing with positions we require the spatial distribution, that is to say, the probability distribution of the centroid with respect to a geometrical origin. We shall calculate the probability distribution of the centroid of the molar system in § 7.

As a guide in the theory, I have found it useful to think of the molar system as a *limit*, which is approached continuously when practical reference objects are gradually changed into a more even and more extensive distribution, without any discontinuity in their mode of employment as reference objects. If we start from the other end, and consider how a distribution in spherical space could be used as a basis for position determinations, the procedure is not so clear.

If σ is the standard deviation of x'_μ, σ represents the uncertainty of position of the molar reference system. It is the counterpart of the uncertainty of *momentum* referred to in (3). Hence, by the uncertainty principle,

$$\sigma \, . \, m_0 c \sim h. \tag{6·1}$$

From the known value of m_0, we obtain $\sigma \sim 10^{-13}$ cm. But the uncertainty principle is not adapted to determine the precise value,* which will be found more directly in § 7.

It should be noticed that, in obtaining ϕ for the molar system, the averaging is over a large but finite number of particles which are in general fairly uniformly distributed. This must not be confused with the averaging over the continuous uniform probability distribution of one particle, which is employed in obtaining expectation values.

The replacement of the N vector particles by N neutral particles with only scalar characteristics enables us to give to each particle a uniform probability distribution; as already explained uniform distribution of a vector is impossible. It also has the effect of reducing the total energy or mass in the ratio 20/136 (*P. and E.* § 14·5). Since the curvature of space is not altered by what is merely a change in the mode of analysis, the neutral particles have effectively a constant of gravitation 136/20 times the usual constant which applies to vector particles and energies. The magnification of energy occurs when the particles in a small volume of neutral matter are assigned wave vectors representing charge and spin—a preliminary step to the introduction of full wave functions which may give them energy additional to the rest energy. It is expensive to transfer energy from the molar to the microscopic system, since the loss to the neutral particles of the molar system will be magnified by the factor 6·8 when they are specified as electrons and protons and so made comparable with the particles of the microscopic system. This is not a violation of the conservation of energy, but a reiteration that the conditions under which current quantum formulae are applicable for the calculation of energy do not allow such a transfer. If such a transfer has been inadvertently made the whole system will appear to have acquired a negative energy.

7—When the material reference system occupies a region whose curvature cannot be neglected, it is necessary to define precisely the mode of computing the centroid. We shall assume tentatively that the particles are to be projected *orthogonally* on to a three-dimensional tangent plane to the spherical space. The centroid is then computed in the usual way. The point of contact of the tangent plane is taken as "geometrical origin", the centroid being the "material origin" from which the observed positions are measured.

---

* A factor $\sqrt{\frac{3}{5}}$ is introduced through m_0 being defined by a limit particle, whereas σ refers to an average particle.

If R is the radius of space, a particle with uniform probability distribution over space has in the three-dimensional projection a root mean square distance $\sqrt{(\frac{3}{4}R^2)}$ from the geometrical origin. The root mean square distance of the centroid of N particles is therefore $\sqrt{(\frac{3}{4}R^2/N)}$; and the standard deviation in one co-ordinate is $\sqrt{(\frac{1}{4}R^2/N)}$. Since N is very large the distribution is Gaussian. Hence the probability that the centroid is in a volume dv at a distance r from the geometrical origin is

$$\pi^{-\frac{3}{2}}k^{-3}e^{-r^2/k^2}\,dv, \tag{7·1}$$

where

$$k = \sqrt{(R^2/2N)}. \tag{7·2}$$

We must justify the use of orthogonal projection in this calculation. If, for example, stereographic projection had been used, k would have been four times as great.

Considering an element of the probability distribution of a particle, the complete stream vector density includes both the momentum vector density and the position vector of the element (*P. and E.* § 5·8). The position vector is the line in four dimensions (or five, if time is involved) from the centre of the hypersphere to the element considered. The essential point is that the process of averaging applied to the wave tensors associates with the average momentum vector an average of the four-dimensional position vectors of the particles. The particles are thereby reduced to a single equivalent particle at their centroid *in four dimensions*. (Protons and electrons have the same weight, since before the averaging they have been replaced by neutral particles.) Thus a treatment, which might perhaps be regarded with some misgiving as too elementary, turns out to be justified.

We can regard the hypersphere as given in the geometrical frame initially, since its radius is determined by the theory in *P. and E.* Chapter XIV. Hence a measure of relative position can be converted into a determination of the relative position vector in four dimensions. By averaging over the N reference particles, we obtain the position vector representing the line joining their centroid to the object particle. This is an observable direction, in contrast to the direction from the centre of the hypersphere to the object particle which is unobservable. These two position vectors define on the hypersphere the material origin and the geometrical origin respectively. The probability distribution of the material origin relatively to the geometrical origin is evidently the three-dimensional orthogonal projection of the probability distribution of the centroid in four dimensions, as calculated in (7·1).

A word of explanation may be offered as to why the procedure of averaging positions and momenta is so different. The difference is due to imposing the condition that the system is in the ground state. The reason for this condition is simply that the current formulae, which we wish to use, presuppose it (§ 2). The condition for a ground state restricts the squares of the momenta to eigenvalues, but imposes no corresponding restriction on positions. We start symmetrically by forming averages in the same way for both, namely by symbolic addition with the E-symbols of primary quantum theory. Then in order to apply to the momenta the condition for a ground state, we transform by second quantization the E-symbolic sum of momenta into a J-symbolic sum with commuting idempotent symbols J. Other differences arise later in the treatment, such as the use of the limit energy m_0 instead of the average $\frac{3}{5}m_0$. But here it is sufficient to notice the origin of the bifurcation, and assure ourselves that for positions the addition remains an E-addition, i.e. vector addition in five dimensions (reduced to four dimensions for the static system here considered).

8—In two dimensions an observed direction is ambiguous to a multiple of 2π; and the mean of N observed directions is ambiguous to a multiple of $2\pi/N$. For $N \sim 10^{79}$, this means that the mean direction is practically indeterminate. There will be a similar indeterminacy of a mean direction in four dimensions, unless by some convention we provide a single-valued reckoning of angles. In § 7 this was furnished by the arbitrarily chosen geometrical origin, which is taken as a pole for the usual single-valued reckoning of spherical polar co-ordinates.* But the geometrical origin (point of contact of the tangent plane) might have been chosen anywhere in the molar system.

Thus two uncertainties are involved in position measures: (*a*) the minute uncertainty found in (7·1), and (*b*) the complete uncertainty of the geometrical origin, which may be anywhere in the material system. This apparent conflict, or at least artificiality, is the result of treating one object particle only, which is too abstract a problem to treat from an observational point of view. When we turn to the more realistic problem of two protons, the two uncertainties sort themselves out; it is then seen that the uncertainty (*a*) refers to the internal particle and (*b*) to the external particle. That the external particle has complete uncertainty of position has already been noticed in § 5; and, in accordance with the representation by double wave functions, ψ_e is fixed by its vector characteristics* in the geometrical frame, whilst ϕ_e gives an even probability distribution of orientation of the molar system—so that the geometrical origin may coincide with any point of the molar system. For the internal particle the potential mobility of the geometrical origin is irrelevant, since the same geometrical origin and plane of projection is used for both protons. It might be thought that the uncertainty (*a*) would also be eliminated; but, as explained in § 6 (3), the protons are supposed to be measured from frames with independent uncertainties, and the two uncertainties combine in the same way as accidental errors.

---

* No origin was provided for the azimuthal angular co-ordinates; but their multi-valuedness does not affect the calculation.

As defined by (5·2) and (5·3) Ψ_i represents the combined distribution of two internal particles. We shall call the system formed by the internal particles of the rest of the universe the *internal molar system*. Such a double internal wave function has already been used in my derivation of the Coulomb energy of a proton and electron (*P. and E.* § 15·6). We may note that the *internal* aspect of the molar system, which focuses attention on the relative positions of pairs of particles, corresponds to the *metrical* aspect of space, which is concerned with intervals between pairs of points. It is therefore the internal molar system which provides a material setting similar, except for its uncertainty, to a geometrical space, and describable by corresponding tensors.

Since $\xi'_\mu = (x'_2)_\mu - (x'_1)_\mu$, and x'_1, x'_2 are uncorrelated, ξ'_μ will have a Gaussian distribution with standard deviation $\sqrt{2}$ times that of $(x'_1)_\mu$. Thus the probability that it falls in a volume dv is

$$\pi^{-\frac{3}{2}} k_0^{-3} e^{-r^2/k_0^2}\, dv, \tag{8·1}$$

where
$$r^2 = \xi_1'^2 + \xi_2'^2 + \xi_3'^2, \quad \text{and} \quad k_0 = k\sqrt{2} = \sqrt{(R^2/N)} \tag{8·2}$$

by (7·2).

The observable quantity is $\xi_\mu - \xi'_\mu$, which is the difference of co-ordinates of the two protons each measured independently from the uncertain material origin. If $\xi_\mu = \xi'_\mu$, we shall say that the two protons are *observationally* coincident; if $\xi_\mu = 0$, they are *geometrically* coincident.

9—We now return to the adjustment required on account of the energy Ω calculated in (4·9). The various eigenstates of the two protons are specified by wave functions which supply a factor modifying their initial uniform uncorrelated distribution (commonly called the *a priori* probability distribution). But the adjustment now to be considered is required before the wave functions are introduced; so that the selected protons still have the initial distribution, and are not in any way different from the unspecified protons. It is the insertion of wave functions which gives them physical prominence by concentrating their probability distribution; the energy is, of course, altered by this last step, but since the change is dealt with by the usual theory of wave functions we are not concerned with it here.

---

* The momentum vector (three-dimensional) of ψ_e is zero, but it will in general possess other vector characteristics, e.g. spin.

Accordingly we consider the N particles all with uniform distribution. We select one of them, a proton, as one member of our microscopic system. For the second particle we may choose either a proton or an electron. The energy assigned to the first proton will jump by $\frac{1}{2}\Omega$ if we substitute the former for the latter choice. This is a matter of assignment of energy only, for the N particles remain distributed just as they were before. Indeed it is unnecessary to choose any particular one of the remaining particles; the jump occurs as soon as we decide to pick a proton, not an electron. The assignment of additional energy to the proton means that there is less energy assigned to the rest of the universe.

The energy contained in a volume element is ultimately a mutual energy of that volume element and the rest of the universe; and its apportionment among the occupants of the volume is a matter of convention. Consider an element dV containing one of the two protons. Besides the selected proton, it will include part of the probability distribution of every other particle in the universe. When we allocate the additional energy $\frac{1}{2}\Omega$ to the proton, the energy allocated to the other particles is correspondingly diminished. The element dV is therefore a singular region of low energy for the other particles.* Considering the Einstein universe of N neutral particles (not yet differentiated into protons and electrons), each particle has a probability dV/V of occupying dV, V being the whole volume of space. Hence, if $-E$ denotes the negative energy of a neutral particle in the field within dV, we have

$$NEdV/V = \tfrac{1}{2}\Omega, \tag{9·1}$$

so that

$$EdV = 2\pi\epsilon^2 R^2/N = 2\pi\epsilon^2 k_0^2, \tag{9·2}$$

by (4·9) and (8·2).

In particular the second proton will (before its selection) have the energy $-E$ for such part of its probability distribution as lies within dV; for, as already noticed, the energy $\frac{1}{2}\Omega$ is transferred before any particular particle is decided on.

---

* Before the transfer of energy $\frac{1}{2}\Omega$, i.e. when the second particle is an electron, the molar system is undisturbed by the microscopic system, and the energy of its particles in dV is the same as elsewhere.

The important question now arises whether the energy $-E$ depends on geometrical or observational coincidence (§ 8) of a particle with dV.* The answer will decide whether $-E$ is associated with a point-singularity in the distribution of the particle or is drawn out into a potential well by probability scattering. We shall find in § 11 that it forms a finite well. Thus the correction to the simple theory of the microscopic system consists in introducing the effect of a potential well around one particle on the energy of the other particle.

10—Before going further, we call attention to the following points:

There is no singularity in the molar system. The singularity dV coincides with the proton; but the proton has uniform probability distribution, so that the reduced energy is spread evenly over the whole system. Nor does it cease to be uniform when we introduce special wave functions for the protons. These introduce correlations, but leave the two protons uniformly distributed (§ 5). The whole effect on the molar system is to reduce the energy of each particle by a constant amount irrespective of its position in space.

It is true that this leaves the energy slightly less than that required for an Einstein universe, so that the molar system, though perfectly symmetrical, is not in self-equilibrium. But the separation into two independent systems is a luxury which simplifies the treatment, not a necessity, and may be deferred. Being unable to separate the systems, we have in § 9 treated the specified and unspecified particles together as one system, and investigated the effect on all the other particles of allocating too much energy to one particle.

The second proton undergoes the same reduction of energy as the other particles, independently of its position in space. But if, instead of position in space, we consider its co-ordinates ξ_μ relative to the first proton, the reduction is a function of ξ_μ. The same would hold for any other particle if we had occasion to consider its ξ_μ.

When we insert a wave function, the initial uniform distribution of the second proton in ξ_μ is replaced by a specialized distribution, which (in the scattering problem) concentrates its probability at points where ξ_μ is small and the corresponding negative energy is large.

In order to compare our results with the current formulae, which treat the microscopic system as completely separate from the molar system, we should, strictly speaking, subtract from the energy of the proton at ξ_μ the small average energy $-\frac{1}{2}\Omega/N$ of its initial uniform distribution, since this energy is needed to bring the molar system to self-equilibrium. But the correction is quite insignificant.

---

* The actual coincidence must evidently be geometrical; but in considering the energy associated with an element of a probability distribution, we have to consider whether the probability distribution is a geometrical or an observational distribution.

The basis of our treatment is that a proton and electron can be separated off to form a microscopic system without disturbing the molar system, so that the particles of the molar system have the same standard energy everywhere irrespective of the positions of the proton and electron; hence when a second proton is substituted for the electron, so that an extra electrical energy Ω is allotted to the microscopic system, the other particles are subject to a deduction of energy in the way we have traced. It is to be remembered that the particles are indistinguishable; and our selected proton consists of proton probability to the amount of one unit withdrawn from the total proton probability in dV—as one might withdraw a pound from a Savings Bank without being able to specify which particular savings it represents. The amount of energy taken out along with the proton is similarly an allocation of a portion of a common store.

The only alternative that suggests itself is that the extra electrical energy $\frac{1}{2}\Omega$ of the proton might be provided at the expense of its own rest mass instead of at the expense of the other particles. A little consideration will, however, show that this is untenable. The rest mass is not a disposable constant that can be adjusted independently for specified and unspecified protons, but has a definite numerical relation to the constant m_0.

11—Coming to the question whether the energy $-E$ depends on geometrical or observational coincidence with dV, we have to examine the use of ξ and $\xi-\xi'$ in current theory, since our conclusion is to be expressed as a correction to current theory. More especially the answer turns on whether the distance r in the Coulomb law refers to ξ or $\xi-\xi'$. The part of the theory chiefly concerned is in *P. and E.* Chapter XV, especially §§ 15·4—15·6, 15·8.

The analysis is based on observed positions $\xi-\xi'$. In particular, the interchange co-ordinate is a *relative* co-ordinate of the object system and comparison fluid. This has been allowed for in *P. and E.*, and the observed distribution has been corrected for the uncertainty of momentum of the comparison system. Momenta referred to the geometrical frame are obtained by multiplying the observable relative momenta by 136/137; the remaining 1/137th gives the energy m_0 representing uncertainty of momentum of the comparison fluid. But there has been no corresponding correction for uncertainty of position, since the investigation in *P. and E.* was not concerned with positions. Accordingly the co-ordinates used in the current wave functions are the uncorrected co-ordinates $\xi-\xi'$; and r in the Coulomb term e^2/r refers to these co-ordinates. This might have been anticipated, because if two particles coincide observationally there is no need to take account of the fact that we are unable to distinguish observationally which is which; whereas if they are observed at two points r apart, the effect of interchanging their identity depends on r and not on their actual distance. In short, if we misidentify the two particles, we interchange the observational errors as well as the actual positions.

The current wave function is therefore of the form $\psi(\xi-\xi')$. We might derive by calculation the corrected wave function $\psi_0(\xi)$. But it is simpler to use the slightly erroneous function $\psi(\xi)$, and apply the necessary corrections to the results. This was done in the application to momenta, the correction giving the factor 136/137. For a geometrical coincidence of position we must take $\psi_0(0)$, which is equal to $\psi(-\xi')$. That is to say, the geometrical coincidence will be represented in the domain of the current wave function by points having the probability scatter ξ'.

This rather inverted way of expressing the result is due to our desire to preserve the current wave function, so that we may use the very complicated calculations of scattering which have been based on it. The essential point is that, since the vanishing of the Coulomb r represents *observational* coincidence, the geometrical coincidence that we are interested in has a probability scatter (about the point $r=0$) equal to the scatter of ξ'.

12—To make up the energy $\frac{1}{2}\Omega$ assigned to the first proton, each uniformly distributed neutral particle contributes an energy $\frac{1}{2}\Omega/N$. In particular the neutral particle, which will on specification become the second proton, has its rest energy reduced from m_0 to $m_0-\frac{1}{2}\Omega/N$. When it is assigned a vector instead of a scalar wave function so as to specify it as a proton, and the adjustment referred to in § 3 is provided for, its rest energy is magnified in the ratio m_p/m_0 determined by the equation

$$10m_p^2-136m_p m_0+m_0^2=0.^*$$

Thus the loss of energy of the second proton (still assuming uniform distribution) is $\frac{1}{2}\Omega m_p/Nm_0$. Since the first and second proton can be interchanged, the total loss of energy of the pair is twice as great, namely

$$\frac{\Omega m_p}{N m_0}. \tag{12·1}$$

We have seen that the loss occurs when $\xi=\xi'$ in the co-ordinates employed in the current wave function. Since ξ has even distribution over the volume V of space, and ξ' has the distribution (8·1), the combined probability is proportional to e^{-r^2/k_0^2}. We therefore set the energy loss at a point distant r from the origin equal to

---

* When a proton and electron are specified simultaneously the magnification is in the ratio $\frac{1}{2}(m_p+m_e)/m_0$ or 136/20, as stated in § 6.

$$Ae^{-r^2/k_0^2}. \tag{12·2}$$

The phrase "energy at a point" refers to the energy term in the wave equation, which is such as to give the energy in a small volume when multiplied by the probability that the particle is in the volume. We can determine A by the condition

$$\int Ae^{-r^2/k_0^2}\,dv/V = \frac{\Omega}{N}\frac{m_p}{m_0}, \tag{12·3}$$

which expresses that the energy at r, multiplied by the probability that the particle is in dv, and integrated over all space, gives the total loss of energy (12·1). From (12·3)

$$\pi^{3/2}k_0^3 A/V = \Omega m_p/Nm_0, \tag{12·4}$$

and, since $V\Omega = 4\pi\epsilon^2 R^2$,

$$A = \frac{4}{\sqrt{\pi}}\frac{m_p}{m_0}\frac{\epsilon^2}{k_0^3}\frac{R^2}{N} = \frac{4}{\sqrt{\pi}}\frac{m_p}{m_0}\frac{\epsilon^2}{k_0} \tag{12·5}$$

by (8·2).

13—The values of N and R are (*P. and E.*, §§ 14·7, 14·9)

$$N = 3{\cdot}145\,.\,10^{79}, \quad R = 1{\cdot}234\,.\,10^{27}\,\text{cm}.$$

Hence by (8·2), $$k_0 = 2{\cdot}20 \times 10^{-13}\,\text{cm}. \tag{13·1}$$

It is convenient to express A in terms of the rest energy $m_e c^2$ of an electron. The observational values give $\epsilon^2/m_e c^2 = 2{\cdot}813 \times 10^{-13}$ cm. By the equation

$$10m_p^2 - 136m_p m_0 + m_0^2 = 0, \quad m_p/m_0 = 13{\cdot}59.$$

Hence by (12·5), $$A = 39{\cdot}2m_e c^2. \tag{13·2}$$

Since I have not paid attention to the factor 136/137, which is liable to occur in the theory, these values may possibly be in error by about one per cent. It should not be very difficult to investigate the factor, and obtain exact values if required.

The observed scattering has been treated by Breit, Condon and Present (1936). Fortunately these authors assumed a Gaussian form of potential well, so that their results are precisely comparable. They give the constants as $k_0 = 2{\cdot}2 \times 10^{-13}$ cm., $A = 39m_e c^2$.

14—For an encounter of two electrons the only change required is the substitution of m_e for m_p in the formulae of § 12. The constant A of the potential well is thereby reduced in the ratio 1 : 1847·6; so that the well potential is relatively much less important, the Coulomb potential being the same.

The investigation provides a foundation for a theory of the nucleus. It may be expected that the extension to more than two particles will encounter mathematical complications of various kinds; but certain broad conclusions may perhaps be permissible:

The extension k_0 of the potential well should not vary much for different nuclei, since it is determined by the uncertainty of the molar system, and has little relation to the object system.

The total intensity Ak_0^3 should be at least approximately proportional to the atomic number. Presumably therefore the depth A will be roughly proportional to the atomic number.

The potential is gravitational not electrical, the energy of particles in the well being proportional to their masses. Thus the well energy of an electron in the nucleus is small compared with its Coulomb energy; and the well energy of a neutron is very nearly the same as that of a proton.

The potential well is determined by the net charge of the nucleus. Thus when well energy is more important than Coulomb energy, it is natural to analyse the system into neutrons and surplus protons; when Coulomb energy is more important than well energy, the natural analysis is into electrons and protons.

In disintegration problems, when a particle leaves the nucleus, it is necessary to define the moment when it ceases to be "specified" and is merged among the unspecified particles; since at that moment there will be a discontinuous change in the adjustment of energy between the microscopic and molar systems.

Corresponding to the present investigation of the effect of unbalanced charge, there should be an investigation of the effect of unbalanced spin; since this will involve an opposite spin of the rest of the universe, making it somewhat different from the ideal Einstein universe presupposed by the current formulae. The spin effect is presumably of much smaller order of magnitude in collision problems; but I think that the neutrino theory is an attempt to deal with it in nuclear phenomena.

The present investigation throws no further light on the structure of *free* neutrons.

SUMMARY

The combination of relativity theory and quantum theory developed in "Relativity Theory of Protons and Electrons" is applied to the problem of close encounters of two protons. It is pointed out that the system of two protons, unlike a neutral system, cannot be treated as simply superposed on the rest of the universe without disturbing it. The disturbance is taken into account, and found to give an extra energy term in the wave equation of amount $-Ae^{-r^2/k_0^2}$, the constants being

$$k_0 = R/\sqrt{N} = 2{\cdot}20 \times 10^{-13}\,\text{cm.},$$

$$A = \frac{4}{\sqrt{\pi}}\frac{m_p}{m_0}\frac{\epsilon^2}{k_0} = 39{\cdot}2 m_e c^2,$$

where R is the Einstein radius and N the number of particles in the universe, m_p the greater root of $10m^2 - 136mm_0 + m_0^2 = 0$, and ϵ the electronic charge. The values agree with those derived from the observed scattering by Breit, Condon and Present (1936).

REFERENCES

Breit, G., Condon, E. U. and Present, R. D. 1936 *Phys. Rev.* **50**, 825–45.
Eddington, A. S. 1923 "Mathematical Theory of Relativity." Camb. Univ. Press.
— 1936 "Relativity Theory of Protons and Electrons." Camb. Univ. Press.

Progressive Lightning

III—The Fine Structure of Return Lightning Strokes

By D. J. Malan,* M.Sc., *Docteur de l'Université de Paris,*

and H. Collens,† M.(S.A.)I.E.E.

(*Communicated by Sir Charles Boys, F.R.S.—Received* 11 *May* 1937)

[Plates 5–10]

1—Introduction

It has been established in two previous papers by Schonland, and the authors (Schonland and Collens 1934; Schonland, Malan and Collens 1935) that the lightning flash involves two consecutive processes, a downward moving leader and an upward moving main or return process. These two processes are repeated for each separate stroke of the series which may make up a complete flash.

In these papers it has been shown that the return stroke travels with a velocity of the order of 2×10^9 cm./sec., and that the velocity decreases as the stroke travels upwards. Variations in luminosity and in velocity have been found to occur after the stroke has passed points where the original leader channel has branched.

The present paper is concerned with a detailed study of these variations in luminosity and velocity, and with their relation to the branches of the channel. It will be shown that the luminosity-time curve at any point along the channel is not a simple one but has an important fine structure (Malan, Schonland and Collens 1935; Malan 1935). This fine structure indicates that at various times after the luminosity is first produced at a point by the upward-moving process, the luminosity rises and falls as the result of the development of additional energy in the channel.

Thus we are led to picture the channel as the seat of a number of "component" return strokes. It will be shown that the evidence indicates that most of these components are called into being by the branches on the channel.

The material available consists of Boys' and other camera photographs of 120 lightning flashes.

---

* The University of Cape Town.

† The Victoria Falls and Transvaal Power Co., Johannesburg.

2—The General Nature of the Return Stroke

In order to simplify the discussion of the evidence as to the fine structure, it is desirable to consider first a general picture of the luminosity-time relation in the return stroke channel. In the case of the first stroke of the series which constitutes a lightning flash, the channel is fairly heavily branched. Fig. 1 is a diagrammatic representation of such a stroke in an idealized vertical channel, as it would be recorded by a camera with one lens moving past a fixed film in the direction indicated by the arrow.

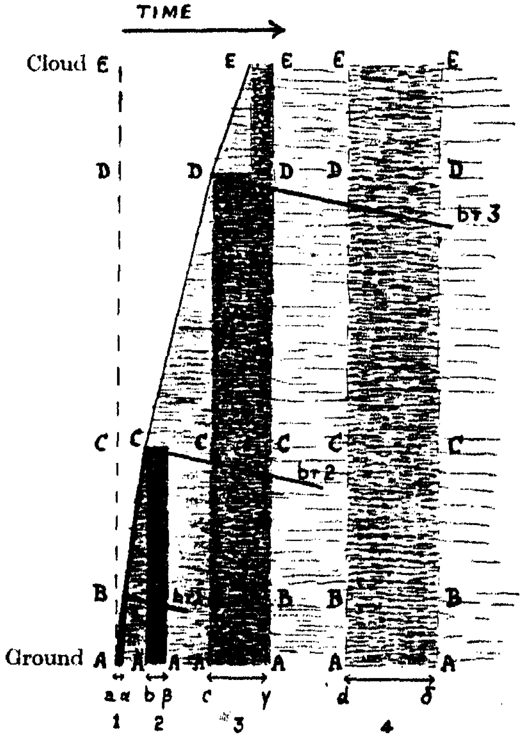


Fig. 1

For any point on the channel, a horizontal line gives the time in the direction of the arrow, and the intensity in terms of the density of the shading in the figure. Thus at the base of the flash the luminosity-time curve indicates the existence of four succeeding components in luminosity, while at the top of the flash these components are two in number, the first being itself double.

These components are distinguished by sharp initial boundaries, *a*, *b*, *c*, *d*, and less clear final boundaries at which their intensity substantially decreases, α, β, γ, δ. The time intervals $a\alpha$, $b\beta$, $c\gamma$, $d\delta$ give therefore values for the effective durations of components 1, 2, 3 and 4, while the time intervals *ab*, *bc*, *cd* give the time separations of the succeeding components. The letters *A*, *B*, *C*, *D*, *E* are used to mark definite points along the channel, while *br* 1, *br* 2, *br* 3 denote successive branches.

The slope of the line *ABCDE* for the leading edge *a* of the first component is due to the lens movement and gives a measure of the velocity of movement of the first appearance of luminosity in the upward-moving return stroke. Its curvature is an indication of the fact that the velocity of edge *a* decreases as component 1 moves upwards.*

Fig. 1 further indicates by the lack of curvature in the leading edges of succeeding components that the measurements to be discussed show that components other than the first travel with a velocity so high that it cannot in general be measured by us. Some of these components, as the diagram shows, are able to catch up with their predecessors, component 2 at branch 2, component 3 at branch 3. (It is probable that an unmarked component 1′ catches up at branch 1.)

The connexion between the components and the branches in the channel will be shown to be very close. Thus component 2 starts at the moment† when component 1 has reached branch 2, and component 3 when component 1 has reached branch 3. Each of these components travels so fast that it extends from the ground to the branch concerned a very short time after it has first been called into existence.

When component 1 reaches *B* (branch 1) it travels outwards along branch 1 and upwards towards *C*. Both the preceding channel length *AB* and branch 1 are observed to be very intensely illuminated, but a definite diminution in intensity is observed in the channel length from *B* to *C* (Schonland, Malan and Collens 1935, p. 618).

At the instant† when component 1 reaches the second branching point *C*, component 2 is observed to start at *A* and to cover the distance *AC* without appreciable delay. This component increases the luminosity along the whole length *AC*, and this increased luminosity is evident also along the branch 2. Again, we observe that the further development of the channel from *C* to *D* proceeds with decreased intensity, as if the channel below *C* were chiefly interested in the development of branch 2. Indeed, if branch 2 is short and there is not a prominent branch which the upward moving fused effort of components 1 and 2 can reach before the completion of branch 2, component 2 ceases and its time of duration $b\beta$ is found to be the same as the time for the whole of branch 2 to be blazed. In case such a prominent branch does exist very near to and above *C*, the structure of component 2 becomes more complicated. Such cases are dealt with in detail in § 7.

---

* Owing to the difficulty of making accurate measurements on short lengths of channel traversed at high velocities, this change in curvature after branching points has only been established for the uppermost portions of a lightning stroke where the velocity has been found to be much lower than that near the base of the stroke.

† "At the moment" or "instant" as used in this connexion in this paper signifies that the events occur within 10μsec. of one another.

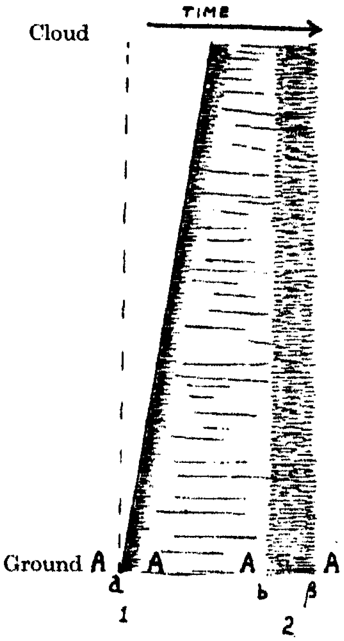


Fig. 2

Component 3 in many cases only lasts until branch 3 is completed. In other examples we find, as in fig. 1, that it continues upwards after branch 3 is completed. This occurs if branch 3 is not exceptionally long and is fairly near the base of the cloud, and suggests that after the completion of branch 3 component 3 is chiefly concerned with a branch or similar call upon its activity hidden in the cloud.

Component 4 starts after the first component has reached the cloud and all the branches have been completed. It is thus concerned with a process taking place in the cloud itself.

There may be also additional components after component 4 which are connected with processes of discharge inside the cloud.

Fig. 2 illustrates what is found in the case of subsequent strokes which carry no branches. In such cases components other than the first start after the first component has reached the cloud and are comparatively weak in intensity.

In cases where a subsequent stroke—generally the second stroke—does carry a branch, it is observed that a component which is connected with the illumination of this branch exists in the same manner as in fig. 1.

The evidence in support of this picture of the fine structure is given in succeeding sections. It must, however, be pointed out that the photographic method of analysing the fine structure is not ideal. The first leading edge *a* of component 1 is clearly marked, but the leading edges *b*, *c*, *d*, etc., of subsequent components are somewhat diffuse and do not lend themselves to measurement to the same degree of accuracy. Further complications arise owing to the tortuosity of the actual channel, which gives rise to difficulty in the observation of the components when the momentary direction of the channel is nearly the same as the direction of relative motion of the lens and the film. In many cases only a few reliable measurements on components other than the first are possible.

For similar reasons, while a great many cases permit of reasonably good investigation upon the original negatives, it is only rarely that an example can be found which will show the fine structure in satisfactory detail for reproduction.

3—METHOD OF MEASUREMENT

The procedure adopted for drawing up the time table for the leading edge *a* of the first component has already been described (Schonland and Collens 1934, p. 655). A contact print is used for these measurements, a fine line being drawn to join corresponding points on the two images of the same stroke formed by the two Boys lenses. The print is cut up and the two corresponding images are mounted side by side so that the two portions of the line are parallel, with the instantaneous direction of motion of the lenses outwardly directed. By measuring the difference in separation between pairs of corresponding points on these images, the time of formation of the leading edge *a* can be determined for any part of the stroke. This edge is considered to start at the base of the stroke (marked *A*) at the zero of time.

As the edges *b*, *c*, etc., of subsequent components are not sharply defined, prints have been found unsatisfactory and the original negatives have been used instead, measurements being made with a transparent glass scale graduated in tenths of a millimetre. The time separation of these edges from the leading edge *a* in the direction of instantaneous motion of the lens is measured at definite points on the stroke. These readings are then added to the times of formation of the corresponding points on the leading edge *a*. In this way a complete time table can be drawn up for all the components of a stroke.

The first leading edge of component 1 can be measured to an accuracy of about 3 μsec. In a few good cases the possible error of measurement for subsequent component edges *b*, *c*, *d*, etc., is 5μsec., but it may rise to 20 or even 50μsec. for weak components of high order whose edges are usually very indistinct.

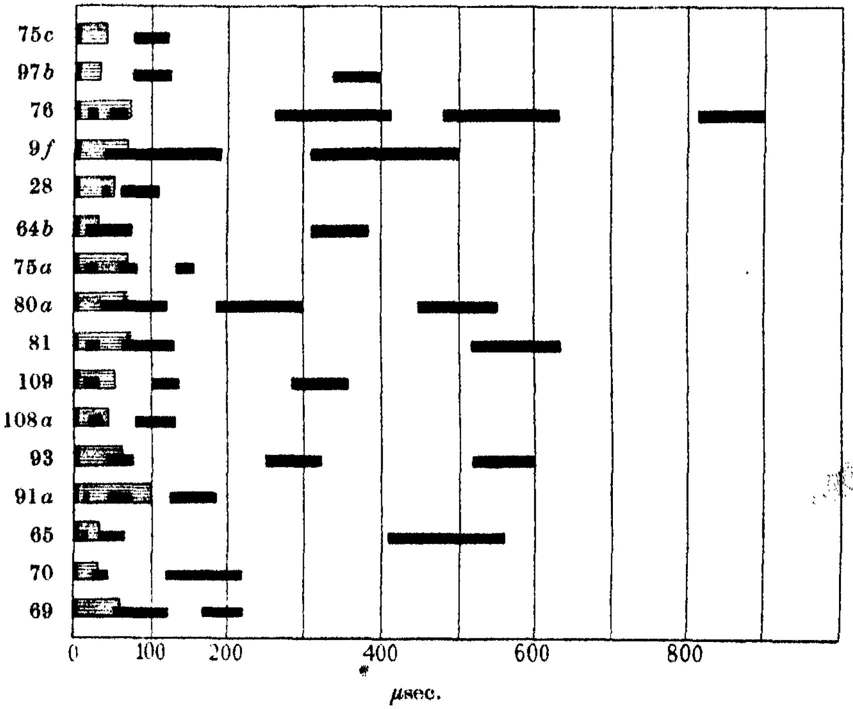


Fig. 3

4—Distribution of Components in Time and Duration

Fig. 3 shows the distribution in time and the duration of the components at the base of fourteen branched and two unbranched strokes, these last being numbers 97*b* and 75*c*. The vertical line at the extreme left indicates the zero of time, which is also the time at which the stroke first starts at the ground (i.e. edge *a* at point *A*, using the terminology of § 2). The figures at the left are the serial numbers allotted to the strokes on the records. The small rectangles shaded with thin horizontal lines indicate the times taken for the stroke to reach the cloud.

The black strips refer to the components making up the complete stroke, the length of the strips indicating their times of duration. The left-hand edges of the strips give the times of commencement and the right-hand edges the times of stopping of the components.

The duration of component 1 at the base of a stroke is very short, and owing to the intense blackening and consequent halation produced on the film at this point, an accurate estimate of the duration is usually impossible. It has been estimated, however, that the upper limit of the duration of component 1 is of the order of 10μsec. This limit is shown by the narrow black strips at the extreme left of the diagram.

Strokes 76, 80 and 81 were too near the camera to get full pictures of these strokes, and the time they took to reach the cloud has been estimated.

It can be seen that for the case of branched strokes the second and sometimes the third components start before the stroke has reached the cloud, whereas strokes 97*b* and 75*c* which are unbranched have their second components starting after the cloud has been reached.

Some of the second components which are apparently of long duration (strokes 69, 64*b*, 9*f*) may be composite (see § 8).

Fig. 4 is a distribution table showing the times at which the different components set out from the base of the discharge. The time of the initial appearance of component 2 is seen to vary from 10 to 50μsec. after the start of component 1, the most frequent value being 25μsec.

The time of appearance of component 3 varies from 30 to 300μsec., while that of component 4 varies from 100 to 500μsec. in extreme cases.

Fig. 5 shows the relation between the durations of components of the same branched stroke. The circles indicate as ordinates the times of duration in microseconds of successive components starting with component 2 at the left. Circles joined by lines indicate components belonging to the same stroke whose serial number is given by the number alongside.

The slope of the lines show that the duration of a component increases with its order, the third lasting longer than the second, the fourth longer than the third, etc. The only exceptions are stroke 69, where the second component lasts longer than the third, and stroke 76, where the fourth and fifth components are of the same duration and the sixth shorter than either of these. The reason for this increased duration is that the prominent

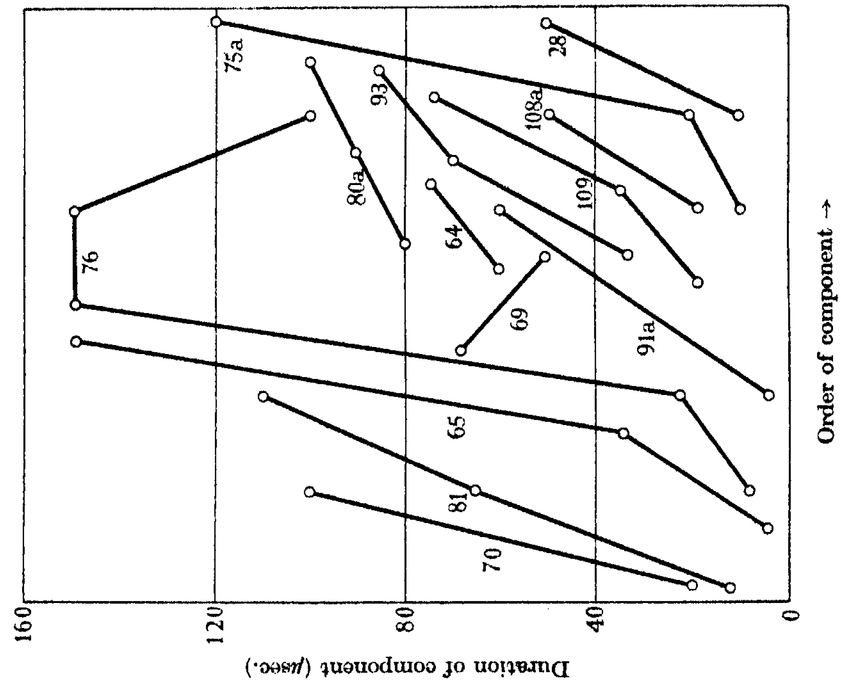

160
120
80
40
0
Duration of component (μsec.)
Order of component →
75a
28
93
108a
80a
109
64
76
69
91a
65
81
70

FIG. 5

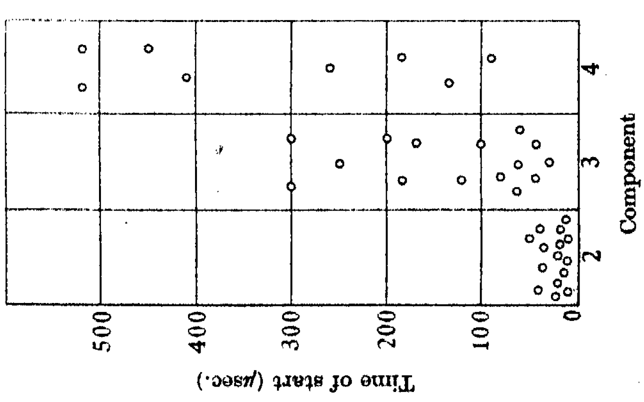

500
400
300
200
100
0
Time of start (μsec.)
2
3
4
Component

FIG. 4

branches high up the track are usually longer and develop more slowly than those lower down.

The most frequent values of the durations and the times of starting of the components of branched strokes are shown in Table I.

TABLE I

Component	Duration μsec.	Time of start μsec.
1	< 10	0
2	20	25
3	50	70
4	100	(150–500)

5—RELATION OF THE COMPONENTS TO THEIR CORRESPONDING BRANCHES

(*a*) *Time of Starting*

It has already been pointed out in the discussion of the general case (§ 2) that the time of starting of a component related to a branch is practically coincident with the time at which the first upward flow of luminosity reaches a branching point in the channel originally formed by the downward moving leader.

In fig. 6, which shows this connexion, the time of starting of a component at the base is plotted against the time of branching of its related branch. Cases of very profuse branching have been omitted, as it is possible that in such cases a component may be related to more than one branch (§ 7), making it impossible to determine which branch was chiefly responsible for the commencement of the component. If the points lay on the line drawn at 45° to the axes, the diagram would indicate that the starting at the base and the branching were simultaneous.

Actually the points tend to lie slightly above this line, so that the time of starting of a component must be considered to lag slightly behind the time of commencement of the branching process. The order of magnitude of this lag is 10^{-5} sec./km. of channel length from the ground to the branch. The possible error of measurement is slightly above 10^{-5} sec./km., so that the actual value of the lag may be less but is certainly not more than this figure.

(*b*) *Durations of Components and their Corresponding Branches*

From the general picture given in § 2 it would be expected that a component would definitely fall off in luminosity as soon as its corresponding branch was completed. This has been established in the case of a few good photographs. Generally, however, precise measurement is impossible on account of the bad definition of the edges β, γ, δ, etc., at which the luminosity decreases. This decrease is much more gradual than the increase in luminosity at the leading edges. In addition, by the time the branch is completed, the initial upward moving luminosity may have passed other subsequent branches which in turn also make a demand on the channel and produce fresh components [§ 7 (*b*)].

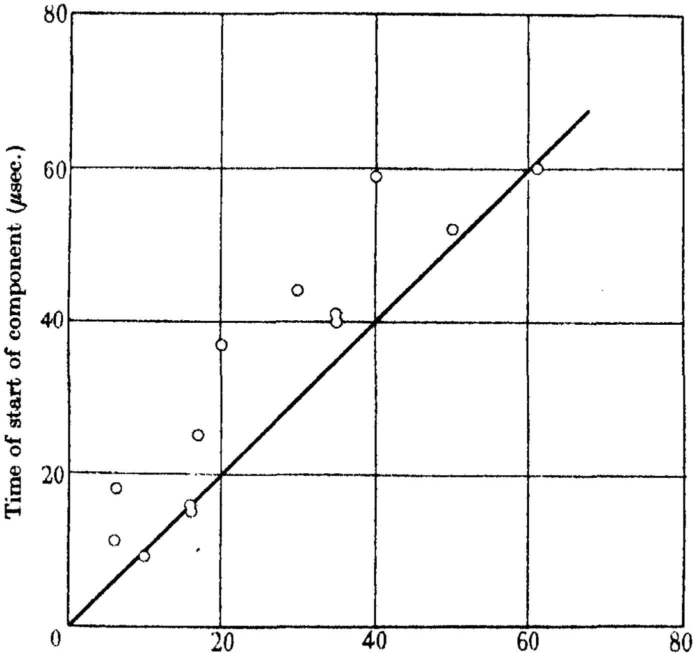


Fig. 6

6—Velocity and Direction of Components

(*a*) *Components Associated with Branches*

It has been mentioned in § 2 that components associated with branches have been found to take a very short time to cover the distance between the ground and the branching point. The magnitude of the time interval will of course depend on the height of the branch above the ground. When this height is relatively small the related component is usually intense and its leading edge is fairly sharp, but the time interval to be measured is extremely short. On the other hand, when the branch is high up the channel, the longer time interval should allow of greater accuracy of measurement, but unfortunately the leading edge in this case is found to be more diffuse than in the former. For these reasons it is in most cases extremely difficult to make an estimate of the velocity, or even to determine the direction of progression of a component. All that can be said as to the direct evidence on this important question is that the direction of movement is indeterminate with our present data, for the velocity in all measurable cases definitely exceeds 10^{10} cm./sec., a value which is the limit of measurement under these conditions.

For components not associated with branches the velocity, as is shown in the next section, is sometimes low enough to be determined.

(*b*) *Components not Associated with Branches*

These components always reach the whole way from the ground to the cloud and are associated both with branched and with unbranched strokes. They constitute the later components of branched strokes and the only subsequent components of unbranched strokes.

An analysis of nine cases where a definite conclusion can be drawn as to their direction shows that two progressed upwards from the ground and seven downwards from the cloud.

Many of these components have very high velocities either upwards or downwards exceeding 10^{10} cm./sec. Others again which have lower velocities are observed to progress downwards only. Four such "cloud" components which were definitely progressing downwards had velocities of $4\cdot10^9$, $2\cdot10^9$, $4\cdot7 . 10^9$ and $2\cdot7 . 10^9$ cm./sec. respectively.

7—Discussion of Specific Cases

Prints of some of the original photographs are given in figs. 7, 13, 15, 16, and 17 (Plates 5–9). They are explained by means of diagrams (figs. 8–12, 14, 18 and 19) which illustrate the different components. The time separations of the components are not drawn to scale on the diagrams, as they would then be too near each other to depict clearly what can be seen on the corresponding photographs.

Definite points along the channel are again marked by the letters *A*, *B*, *C*, etc. A figure next to one of these letters at the first leading edge denotes the time in microseconds at which the return stroke first reaches that point. The zero of time is again the time of starting of the stroke at the ground. A figure at a point next to a branch indicates the time the luminosity first reaches that point. Thick arrows show the direction of instantaneous motion of the lens of the Boys camera.

Tables are also given of the time in microseconds when the edges b, β; c, γ; d, δ, etc., reach the points A, B, C, etc.

(*a*) *Strokes with Single Branches*

(i) *Stroke* 70.

TABLE II (figs. 7 and 8)

(Time in microseconds after start of first component at A)

Point on track	Comp. 1	Comp. 2		Comp. 3	
	a	b	β	c	γ
A	0	20	44	120	220
B	3	20	40	—	—
C	5	17	47	115	—
D	10	17	46	—	—
E	17	—	—	—	—
F	16	(P 28)	—	116	—
G	26	—	—	110	—
H	28	—	—	—	—

Table II shows that component 2 starts at A (fig. 8) shortly after component 1 reaches the branching point above E. Component 2 is interested in the branch until it is completed at a little over 23μsec. By this time the first component has reached the cloud and some hidden branching process is probably responsible for the fact that component 2 progresses upwards to the cloud as shown by the shaded strip at P (fig. 8) which starts at 28μsec.

Component 3 is interested in the cloud only and starts at about 90μsec. after component 1 has reached the cloud.

(ii) *Stroke* 64*b*.

TABLE III (figs. 7 and 9)

(Time in microseconds after start of first component at A)

Point on track	Comp. 1	Comp. 2		Comp. 3	
	a	b	β	c	γ
A	0	15	74	310	384
B	5	—	—	—	—
C	16	—	68	—	—
D	22	—	60	—	—
E	32	—	—	272	—

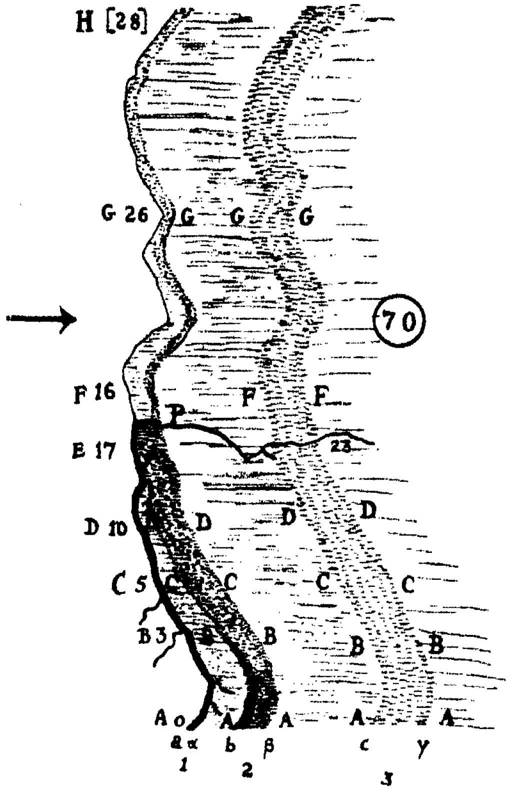

H [28]
G 26
G
G
G
70
F 16
P
F
F
23
E 17
D 10
D
D
D
C 5
C
C
C
B 3
B
B
B
B
A o
A
A
A
A
a α
b
β
c
γ
1
2
3

FIG. 8

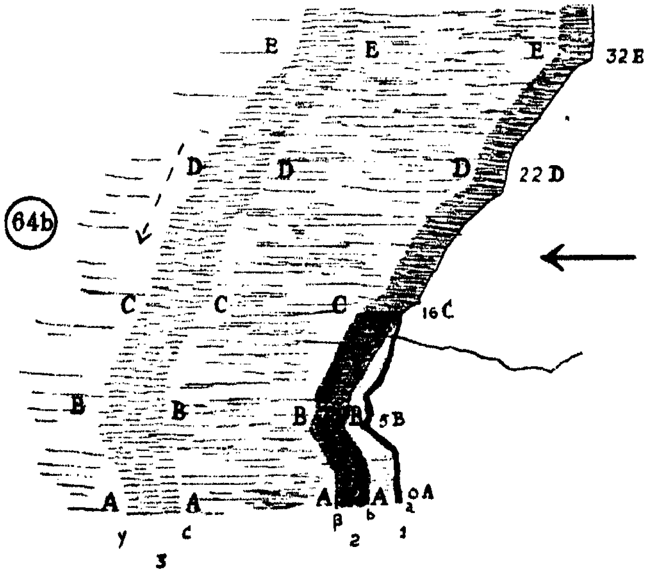

E
E
E
32 E
D
D
D
22 D
64b
C
C
C
16 C
B
B
B
B
5 B
A
A
A
A
o A
a
γ
c
β
b
2
1
3

FIG. 9

Here component 2 again starts when the first upward flow of luminosity reaches the branching point at *C* (fig. 9), namely 16μsec. after component 1 has started at the ground. It catches up component 1 at *C* and lasts till 74μsec., i.e. 30μsec. after component 1 has reached the cloud.

Component 3 is interested in the cloud only and definitely moves downwards (indicated by the dotted arrow in fig. 9) with a velocity of $4{\cdot}7 \times 10^9$ cm./sec. Branch *C* is very faint on the print and has been indicated by a series of white dots.

(*b*) *Heavily Branched Strokes*

(i) *Stroke* 65.

TABLE IV (figs. 7 and 10)

(Time in microseconds after start of first component at *A*)

Point on track	Comp. 1	Comp. 2		Comp. 3		Comp. 4	
	a	*b*	β	*c*	γ	*d*	δ
A	0	9	13	30	65	410	560
B	2	—	13	—	—	—	—
C	4	—	—	26	63	374	—
E	11	—	16	29	—	—	—
F	12	—	—	—	71	348	—
G	18	—	—	—	92	—	—
H	26	—	—	—	—	[326]	—

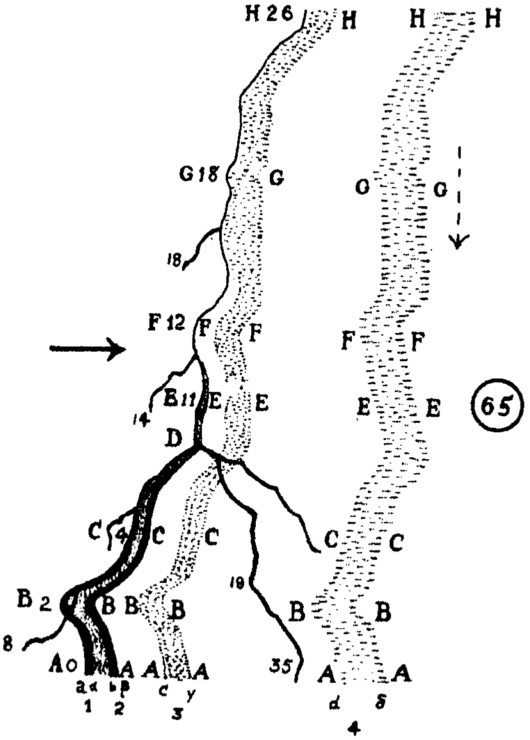


FIG. 10

Flash 65 consists of a single heavily branched stroke. Component 2 starts at the ground when component 1 reaches the branching point at D, viz. at about 9μsec. after the stroke has started at the ground. Above D, component 2 is also interested in the branch above E as illustrated in the diagram. Component 3 starts at about 30μsec., i.e. when component 1 has reached the cloud.

In component 4 the luminosity travels downwards (indicated by the dotted arrow) with a velocity of 4×10^9 cm./sec.

The next few illustrations (figs. 11–14) show how more than one component can contribute to the completion of a branch.

(ii) *Stroke* 93.

TABLE V (figs. 7 and 11)

(Time in microseconds after start of first component at A)

Point on track	Comp. 1	Comp. 2		Comp. 3		Comp. 4	
	a	b	β	c	γ	d	δ
A	0	41	[74]	250	320	520	600
B	7	—	—	—	—	—	—
C	9	40	—	—	—	—	—
D	20	—	—	260	—	—	—
E	27	49	—	—	—	—	—
F	33	—	92	253	—	—	—
H	47	—	—	—	—	—	—
L	60	—	—	—	—	—	—

Component 1 is probably double, its second part being initiated by the branch at C, but owing to the intense blackening on the film its two parts cannot be distinguished.

It will be seen from fig. 11 that the component marked 2 is first called into being by the branch at G. Before the latter is completed, component 1 has reached the branches at H and at K which are not very far beyond G. These branches then start developing slowly as shown by the "drag" on the moving lens record (fig. 11). The main channel is in the first place chiefly occupied in the completion of branch G. When this is effected, the channel directs its attention to branch H so that component 2 does not lose its luminosity at this stage. The same process is again repeated for the branch at K after the completion of the one at H. In the meantime the stroke has reached the cloud, where a hidden branching or other process (at 80μsec.) makes yet another demand on the channel below and causes component 2 to progress right up into the cloud before the long branch K

is fully completed. This is shown in fig. 11 by the shaded band above *K* immediately to the right of component 1.

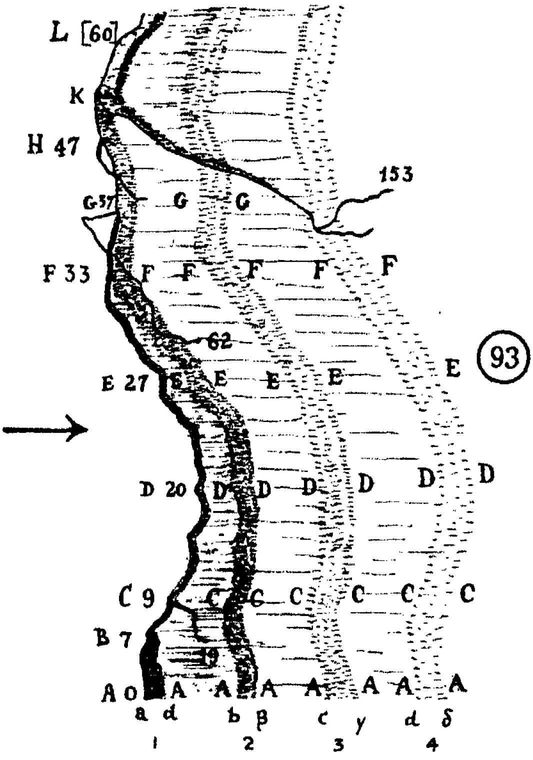


Fig. 11

The result is that component 2 lasts much longer than the 25μsec. required for the completion of the branch from *G*. It is in fact related to the three branches *G*, *H*, *K*, and to an additional branching process in the cloud and has a total duration of the order of 50μsec. Components 3 and 4 are interested in the cloud only

(iii) *Stroke* 75*a*.

Stroke 75*a* is the first heavily branched stroke of a series of seven strokes. Component 2 starts when component 1 reaches the branching point at *D* at 15μsec. (fig. 12). Before branch *D* is completed, component 1 has reached the two branching points at *F* and the two branches, *br* 4 and *br* 5, start developing slowly. As soon as branch *D* is completed, the lower branch at *F* (*br* 4) receives the full attention of the channel to ground and

TABLE VI (figs. 7 and 12)

(Time in microseconds after start of first component at *A*)

Point on track	Comp. 1	Comp. 2		Comp. 3		Comp. 4	
	a	*b*	*β*	*c*	*γ*	*d*	*δ*
A	0	16	[36]	60	80	135	157
B	2	—	—	—	—	—	—
C	7	—	—	—	—	—	—
C'	—	—	—	60	—	125	—
D	15	—	—	—	89	—	155
E	21	—	—	—	—	—	—
E'	—	—	—	68	—	140	—
F	32	—	—	—	—	—	152
F'	—	—	—	64	95	—	—
G	43	—	—	58	—	133	153
H	61	—	—	—	—	—	—
J	64	—	—	—	—	—	164

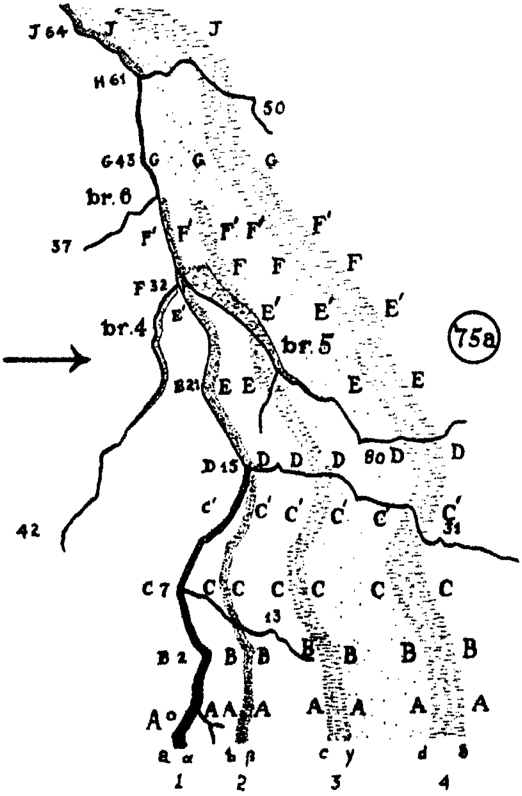


FIG. 12

is completed to its tip. By this time (42μsec.) the next branch (*br* 5) has progressed an appreciable distance, the branch above F' (*br* 6) has been fully completed and the main channel luminosity has reached the point G. Branch 5 now being far advanced does not seem to be able to make a heavy demand on the channel and slowly advances (shown by the "drag") while component 2 decreases in intensity. In the meantime, however, the point H is reached where the last visible branch causes the initiation of component 3 which then also helps in the final completion of branch 5 as is clearly shown in fig. 12.

Component 4 is concerned with events taking place in the cloud.

(iv) *Stroke* 91*a*.

TABLE VII (figs. 13 and 14)

(Time in microseconds after start of first component at A)

Point on track	Comp. 1	Comp. 2		Comp. 3		Comp. 4		*
	a	b	β	c	γ	d	δ	δ'
A	0	11	15	44	75	—	—	220
B	1	8	12	—	—	—	—	—
C	11	—	—	—	—	—	—	—
D	17	—	—	39	—	107	—	—
E	18	—	—	—	—	—	166	—
F	26	—	—	—	—	100	—	—
G	47	—	—	—	—	—	157	—
H	48	—	—	—	—	85	—	—
K	[98]	—	—	—	—	—	—	—

Fig. 13, Plate 6, shows the original photograph (inset) and an enlargement of the base of stroke 91*a*. (Flash 91 consisted of a single stroke to ground.)

Component 2 is initiated by the branch at C. While this branch is being completed, the luminosity progresses faintly higher up the channel and along the longer branch R immediately above C. As soon as the whole of the branch C is blazed, the branch R develops more energetically, increasing its luminosity and that of the main channel between C and R. This is shown by the shading in fig. 14 and on the photograph in fig. 13 where the leading edge of the channel of the first component between C and R is picked out with three white dots and that of the second component with three black dots.

Component 3 is called into being by the long branch above F and is also responsible for blazing the short bright branch at G immediately beyond it.

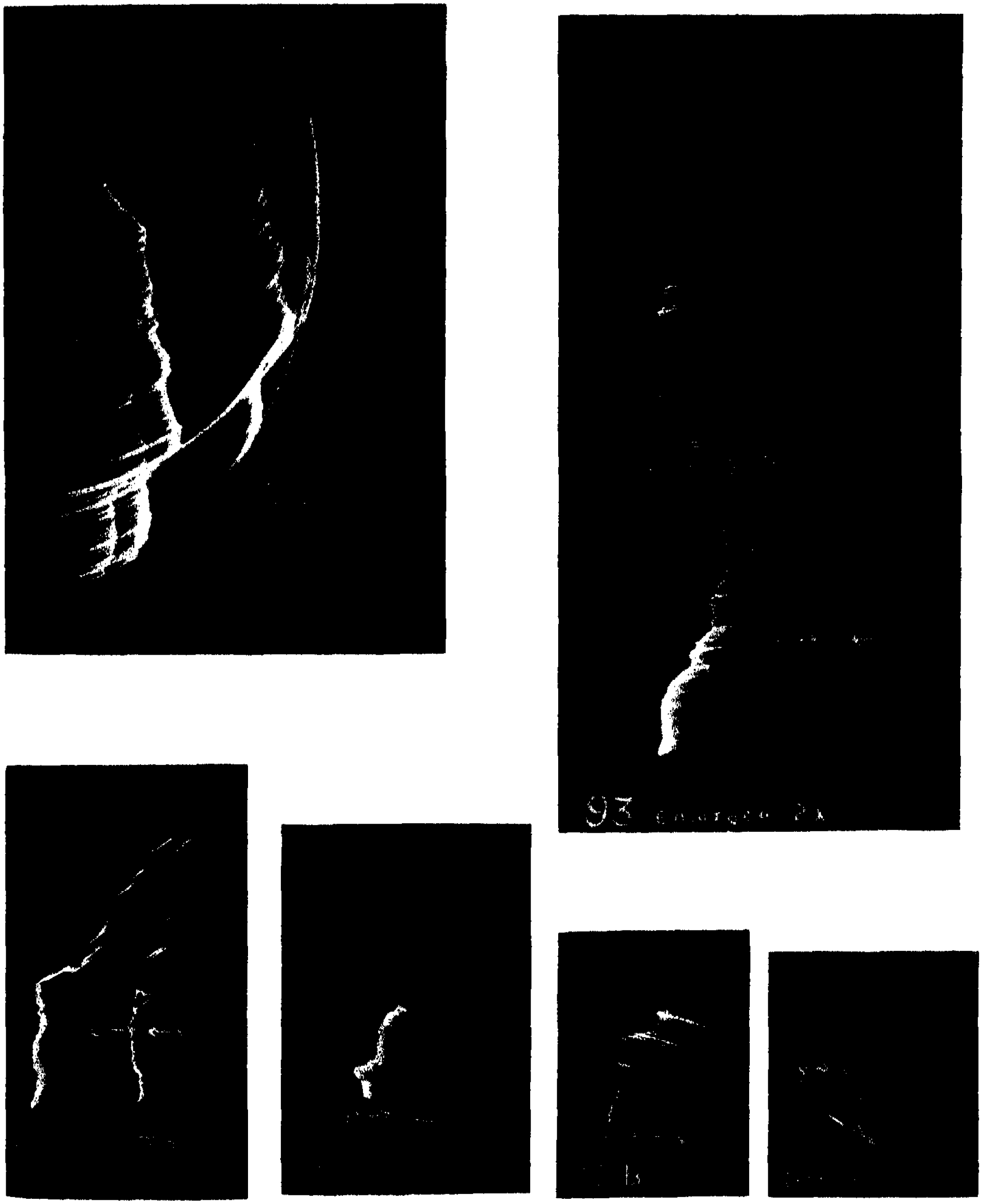


FIG. 7

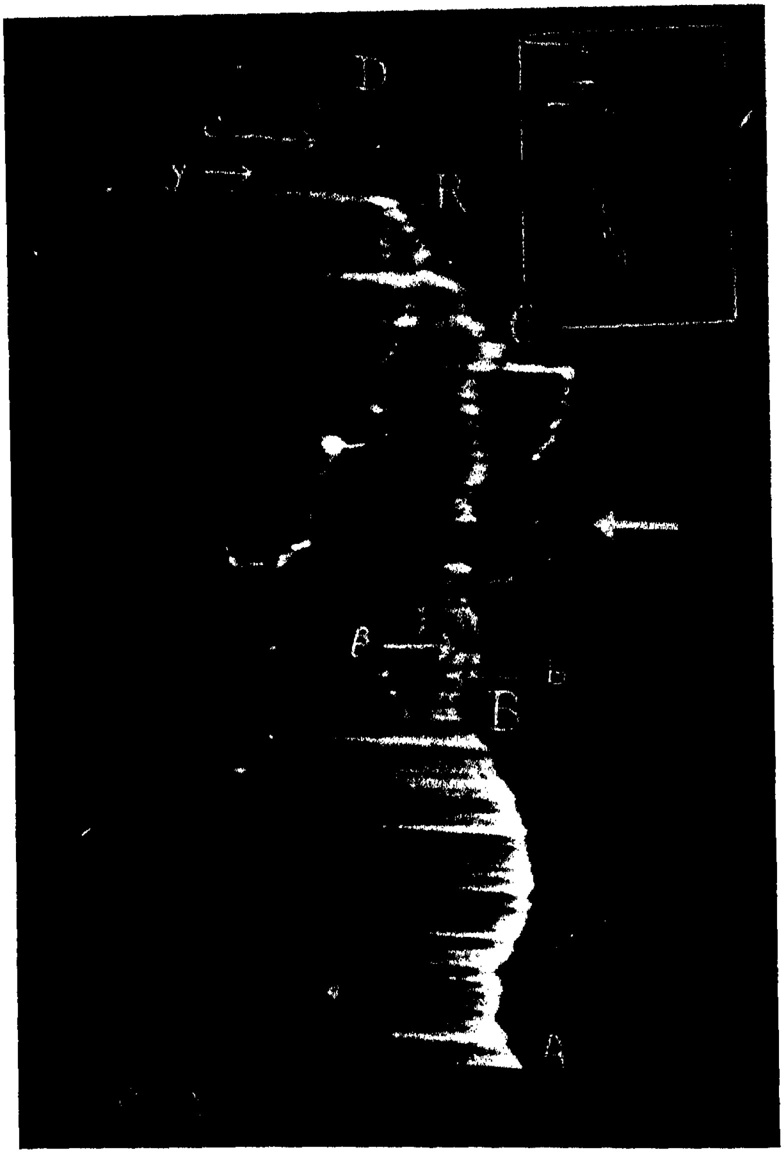


Fig. 13

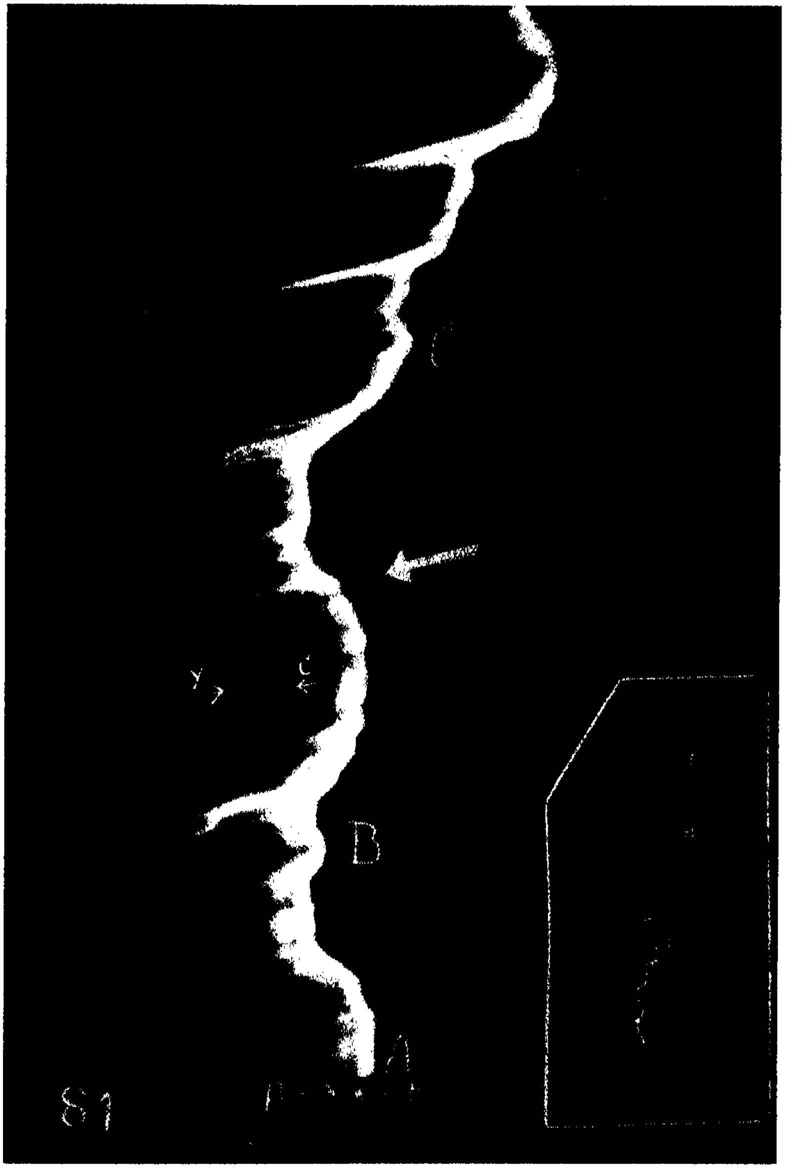


FIG. 15

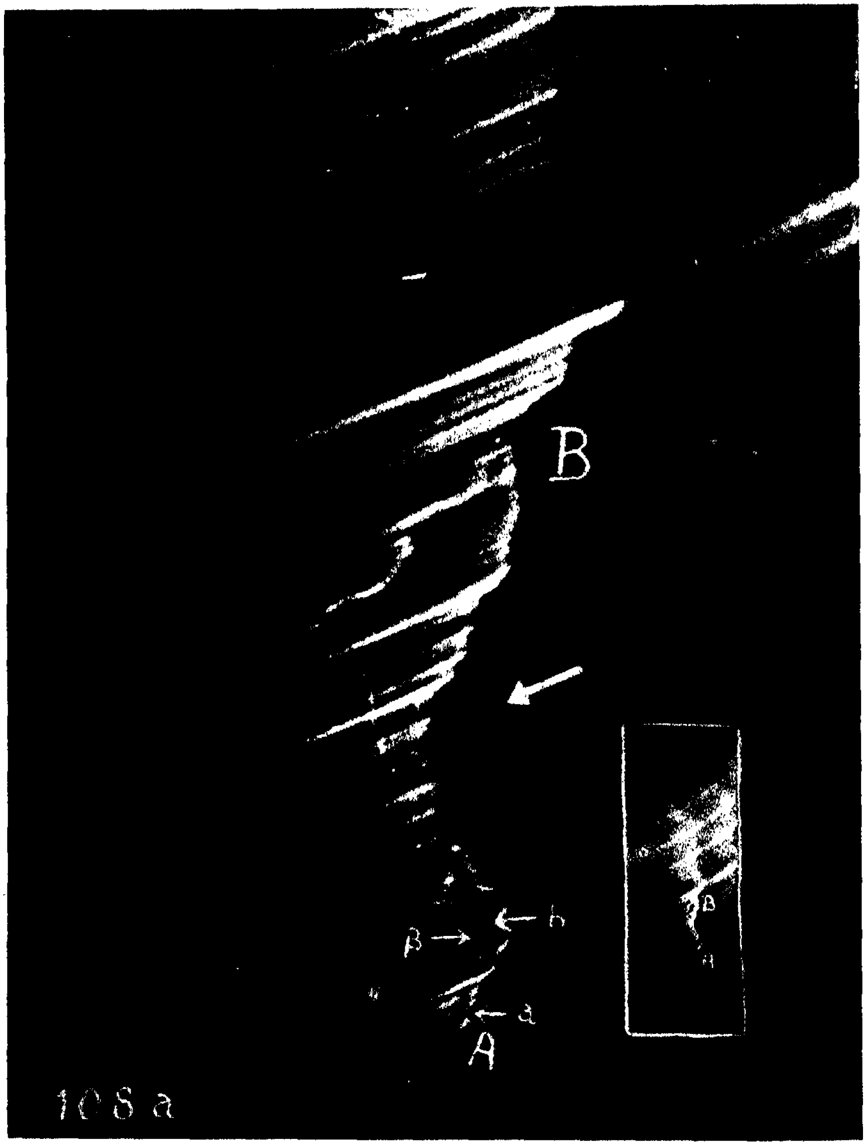


Fig. 16

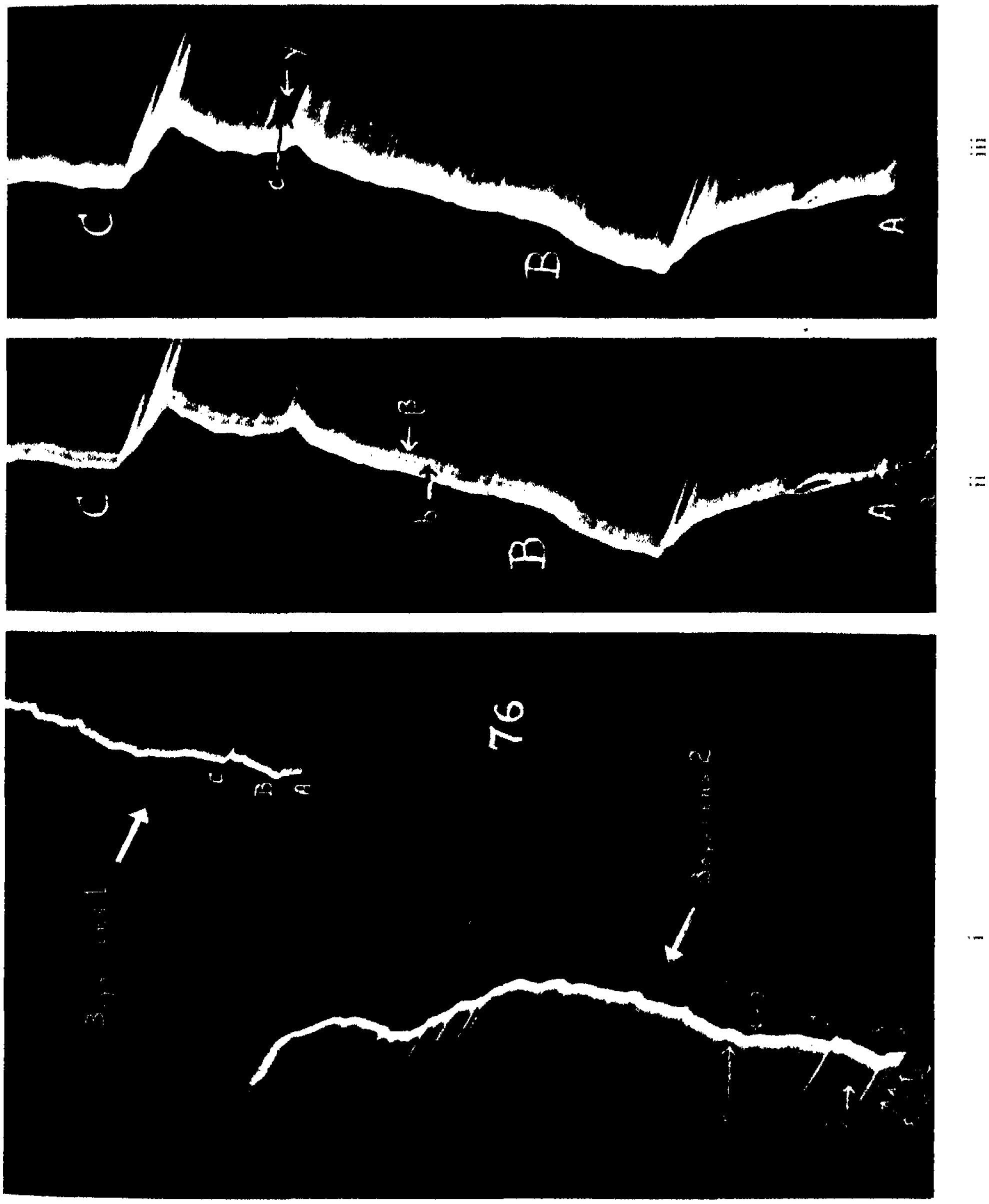


FIG. 17

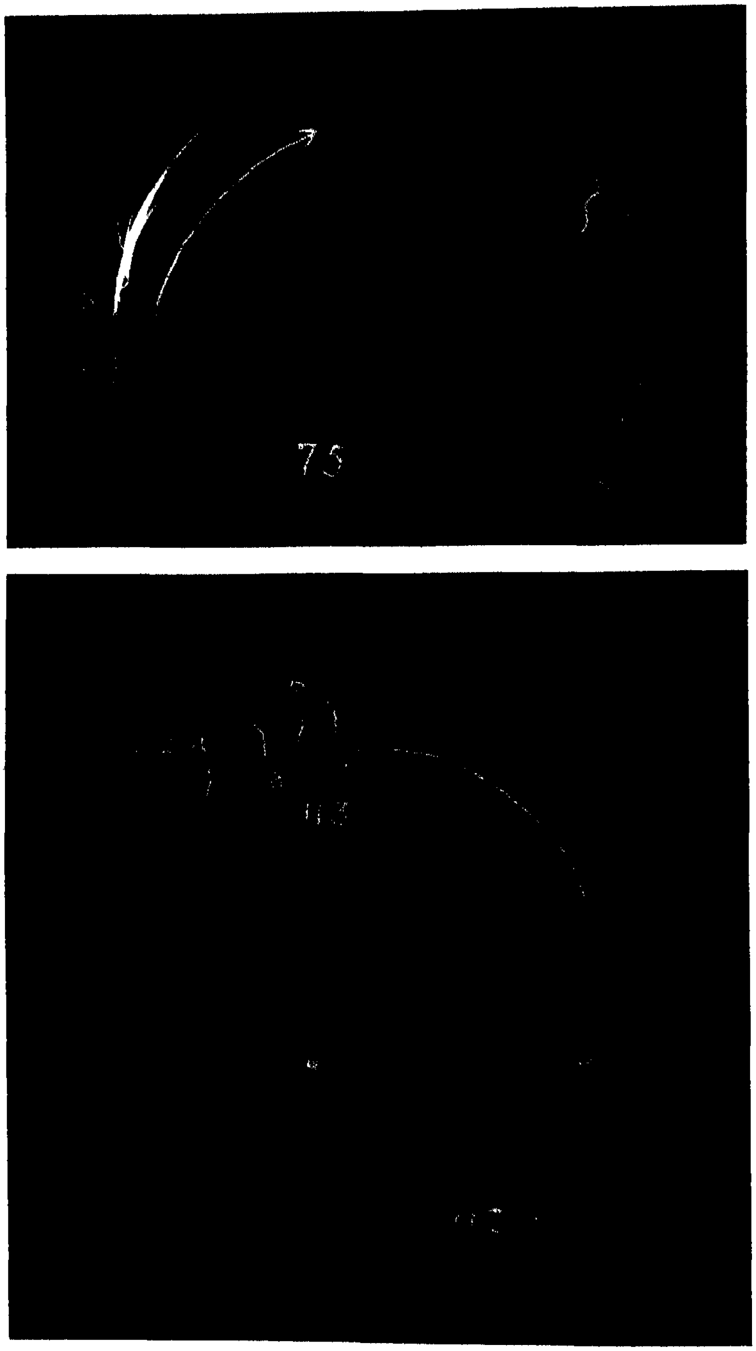


Fig. 20

Component 4 is double, its dual nature only being distinguishable above *F*. It aids in the final completion of the branch at *S*. It also has a faint continuation until the time marked by δ' (fig. 14).

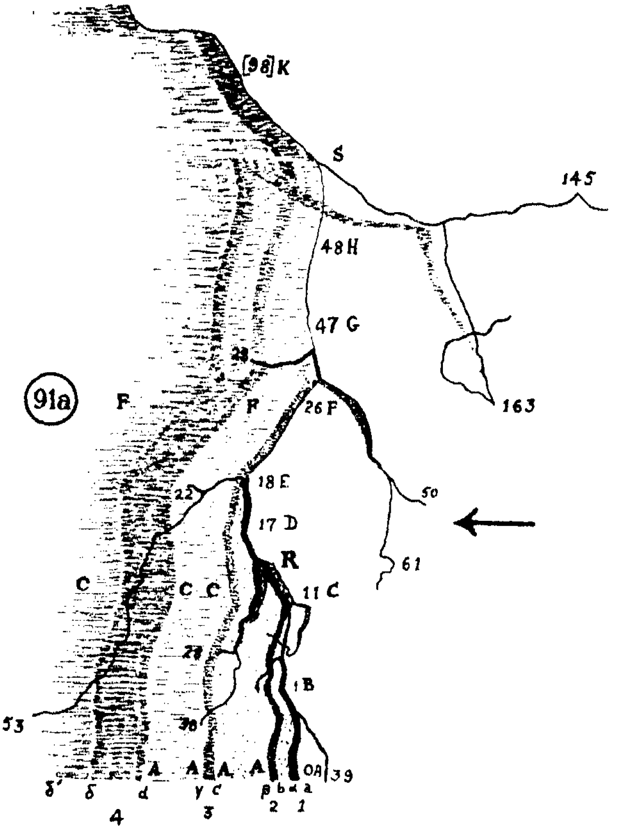


FIG. 14

The next two figures (15 and 16, Plates 7 and 8) are photographic enlargements showing component 2 more clearly than is usual. These strokes (81 and 108*a*) have components following the general behaviour described above, so that it would be superfluous to describe them in detail by means of diagrams.

(v) *Stroke* 81 (fig. 15, Plate 7).

This gives the original photograph of the whole track and an enlargement of the lower portion *ABC* of stroke 81 showing the edges *b* and β

of component 2 exceptionally clearly. Component 3, edges c and γ, have also been marked.

(vi) *Stroke* 108*a* (fig. 16, Plate 8).

This shows an enlargement of the portion AB of stroke 108*a* (inset). Component 2 edges b and β can be clearly distinguished.

(vii) *Stroke* 76 (fig. 17, Plate 9).

Fig. 17 (i) shows the complete Boys camera record of stroke 76.

Flash 76 consisted of a single stroke to ground at a distance of 2 ± 0·5 km., being so near to the camera that only the lower portion of the stroke could be photographed. Lens 1 being near the top of the film, recorded the lower portion of the stroke up to the first branch only. (Lens 2 was left at the full aperture of $f/6{\cdot}3$ and lens 1 was stopped down to $f/16$ for the purpose explained below in § 10*b*).

Fig. 17 (ii) and (iii) are enlargements of the portion ABC of the Boys lens 1 picture. (ii) has been printed dark to show up components 1 and 2 more clearly. They can be seen separated by a narrow dark line (of width αb) running up the track. αb represents the time separation between these components. (iii), which is printed more lightly than (ii), brings out component 3 with its leading and finishing edges c and γ respectively.

On the Boys lens 2 picture of (i), components 1, 2 and 3 are lost in the intense white line (they can be seen on the negative), but the leading edge d of component 4 and the finishing edge ϵ of component 5 can be clearly distinguished. The positions of the other edges (δ, e) of these components are indicated by arrows. A faint sixth component also exists.

A time table of the start and finish of the various components at the base of the flash is given below:

TABLE VIII (stroke 76, fig. 17)

(Time in microseconds after start of first component at A)

Point on track	Comp. 1		Comp. 2		Comp. 3		Comp. 4		Comp. 5		Comp. 6	
	a	α	b	β	c	γ	d	δ	e	ϵ	f	ϕ
A	0	11	18	26	46	67	260	410	480	630	815	900

(c) *Unbranched Strokes*

The results of an analysis of two typical examples of unbranched strokes (75*c* and 97*b*, fig. 7, Plate 5) are shown in figs. 18 and 19. Full time tables of all the components are given on the diagrams.

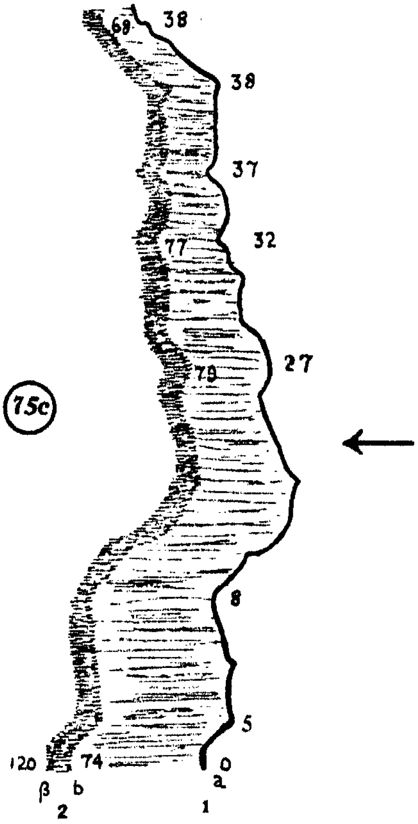

β
b
a

FIG. 18

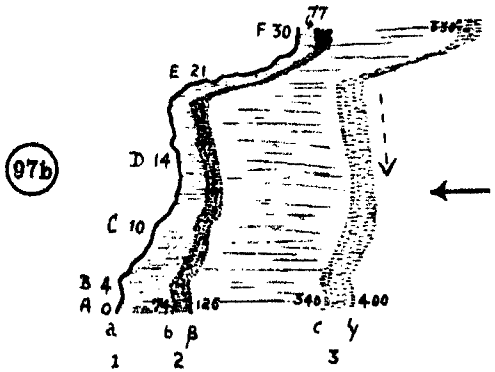

97b
77
F 30
E 21
D 14
C 10
B 4
A 0
126
340
400
a
b
β
c
γ
1
2
3

FIG. 19

It will be seen that all the subsequent components start after component 1 has reached the cloud. Component 3 of stroke 97*b* progresses downwards as indicated by the dotted arrow. The edges of component 2 of this stroke and component 2 of stroke 75*c* are too uncertain to state definitely in which direction they progress.

8—THE TOTAL NUMBER OF COMPONENTS

Owing to the complexity of the fine structure which has been described and the difficulties of photography, it is safe to surmise that in several of the examples cited in this paper the total number of components is higher than that which has been observed.

Referring to figs. 11, 12, 14, it is probable that the very intense parts at the bases of these strokes up to the first branch really consist of two components. The over-exposure and the small time separation of these components would not allow them to be resolved. An examination of photographs taken with a camera of higher time resolution, having a film mounted on a revolving drum (Schonland, Malan and Collens 1935; Boys 1929), shows such cases to exist. These photographs are, however, too diffuse for reproduction.

It can thus be expected that each branch will have its own component unless, as described in § 7, it is so situated as to be able to make use of a component initiated by another branch lower down the channel. The maximum number of components that has been observed for a first branched stroke is 6, the most frequent number being 4. Unbranched strokes rarely have more than two and mostly have only a first component. In exceptional cases the number may even go up to 20 (§ 9, stroke 113*b*).

9—TOTAL DURATION OF CHANNEL LUMINOSITY

The luminosity along the main channel may persist for a relatively long time (Schonland and Collens 1934; Walter 1935; McEachron and McMorris 1936), and a duration as long as 0·23 sec. has been mentioned in the literature (McEachron and McMorris 1936).

The total duration of channel luminosity as measured on the photographs taken with a Boys camera would of course depend on the conditions of photography. A specific instance is that of stroke 76, fig. 17 (Plate 9), where the apertures of the two lenses were *f*/6·3 and *f*/16 respectively. The *f*/16 lens photograph shows a total duration of 2200μsec., while the $f/6{\cdot}3$ photograph gives a duration of 4000μsec. owing to the greater light-gathering power of the lens.

The most frequent value of the duration as measured on the photographs appears to be of the order of 1000μsec., ranging in extreme cases from a few hundred microseconds to half a second.

The long "drag" (shown by the arrows) on the slow-moving camera pictures of strokes 75*b* and 113*b* (fig. 20, Plate 10) shows these strokes to be of exceptionally long duration, lasting for 0·25 and 0·50 sec. respectively. The property of long duration seems to have no relation either to the order of the stroke (McEachron and McMorris 1936) or to its relative intensity as compared to the other strokes composing the flash.

Stroke 75*b* shows three faint components in the "drag" following the first intense component, and stroke 113*b* actually shows at least twenty such components.

The suggestion of Walter (1935) that the long duration is due to a high ground resistance has already been refuted by McEachron and McMorris (1936), who pointed out that the different strokes of a flash to a steel structure showed different times of total duration, although the ground resistance was small and was the same for all the strokes.

It has been pointed out in this paper that the earlier components of branched strokes are related to the branches irrespective of the nature of the ground, and that a later component sometimes starts immediately after the stroke has reached the cloud, so that in all probability the latter is related to a branching or other process taking place in the cloud. From the general nature of these components it will be legitimate to assume that still later components, which are similar in appearance and behaviour to the former, are also related to processes in the cloud itself. The existence of several components in most strokes of long duration thus strengthens the view of McEachron and McMorris that long duration is entirely due to processes in the cloud and has no relation to ground resistance.

10—INTENSITY DISTRIBUTION

(*a*) *Variation Along the Channel*

It is known that there is a distinct diminution in intensity of component 1 after passing points of branching (Schonland, Malan and Collens 1935, p. 618). A discussion of two specific examples will serve to show the marked nature of these variations.

Referring to stroke 65, fig. 10, it will be seen that component 1 is very bright up to the branch at *C*. After reaching *C* it is chiefly concerned with this branch which is consequently very intense. Component 1, however, continues progressing upwards along the main channel with greatly diminished luminosity till it reaches the branching point of the long branch at *D* a short distance above *C*. Component 2 is now called into being as a result of the development of this branch. From *C* to *D* component 2 now actually becomes brighter than the original component has been along this part of the channel. From *D* to *E* components 1 and 2 have fused after a marked diminution in intensity of the latter at *D*. Component 3 catches up with the first leading edge just above *F* with the result that from here to the cloud the channel is illuminated at any point along its path by a glow of low intensity lasting for about 35μsec. after it has first become visible.

Stroke 91*a* (figs. 13 and 14) shows a similar effect. From *A* to *C* component 1 is the brighter of the first two components but loses most of its intensity beyond this point. In the short distance from *C* to *R*, where the next branch is situated, component 2 becomes brighter than its predecessor, as is clearly shown both in figs. 13 and 14. From here to *E* they have merged and are of the same intensity. Beyond *E* they have again fused with component 3, giving a much fainter illumination lasting for about 30μsec. at any point from *E* to *F*.

A frequent type of variation in channel luminosity illustrated by the foregoing is as follows:

We may suppose a short branch to be reached by the first component of the main stroke—which may or may not have fused with the second component. This short branch causes a fresh component to run along the channel, catching up the former and delivering energy to the branch. At the same time the first component is travelling upwards with a reduced luminosity, and if it reaches a long branch a short distance farther on, this branch will start developing faintly. It will not have progressed very far, however, before the former branch is completed, and since the channel will now be free to supply it with energy, it will make a demand on the last component, causing the illumination of the channel to increase till it may surpass its former intensity along its track between the two branches under discussion.

If the upward moving luminosity reaches a branch after all the previous branches have been totally or almost totally completed, this branch will call into being a component of its own.

(*b*) *Variation with Time at any Point Along the Track*

A luminosity-time curve at the base of any stroke will have as many maxima as the total number of existing components of the stroke. This will be true for any point along the channel for *unbranched* strokes. If the stroke is branched, however, the number of components may become less after the passage of branching points according to the general principles discussed in § 10*a*. Thus the stroke illustrated by fig. 1 has four components from A to C, three from C to E and only two above E.

A detailed photometric study of stroke 76 (fig. 17), made in order to determine the relative intensities of these components, will now be discussed. In taking photographs with the lenses of the Boys camera at full aperture ($f/6{\cdot}3$) it is found that owing to the intensity of the luminosity of the first component it is often over-exposed on the film. To obtain a normally exposed image of this component it was found necessary to use a smaller aperture on one of the lenses ($f/16$, § 7 (*b*) (vii)). Some of the weaker components recorded by this lens will be under-exposed but may have a normal exposure in the full aperture picture.

Knowing the apertures of the lenses in the two cases, it is possible by the application of the principles of photographic photometry to find which parts of the images are normally exposed and also to compare the light-intensities of the different components.

A recording microphotometer (Lambert and Chalonge 1926), where the image of an illuminated slit is focused on the photographic film and the intensity of the transmitted light is measured by means of a photoelectric cell, has been used for this purpose. Since the lenses of the Boys camera move in circles, the film is mounted on a special table having a circular motion.

Fig. 21 shows a curve obtained in this manner at the point C on the negative of the Boys lens 1 image of stroke 76 (fig. 17, Plate 9).

The zero line AB represents the photoelectric current when no light falls on the photocell. The illuminated slit is made to move from right to left across the negative of the upper image of fig. 17 (i), the recording of fig. 21 being taken from left to right. The slit is first obscured so that a small portion of the zero line can be traced. When the obstruction is removed the light passes through the clear part of the film to the right of the photographic image and the curve moves upwards off the paper. As the slit image approaches the more intense part of the film the curve moves down towards the right till it gives the minimum for component 3 whose edges are indicated by γ and c. When it successively crosses components 2 and 1, two further minima are traced out but this time much lower down owing to the greater blackening produced by these components whose edges are marked by β, b and α, a respectively. After passing the edge a the light traverses the clear part of the film and the curve again moves upwards off the paper. To get the zero line the slit is again obscured. This is done because the axes are not rectangular.

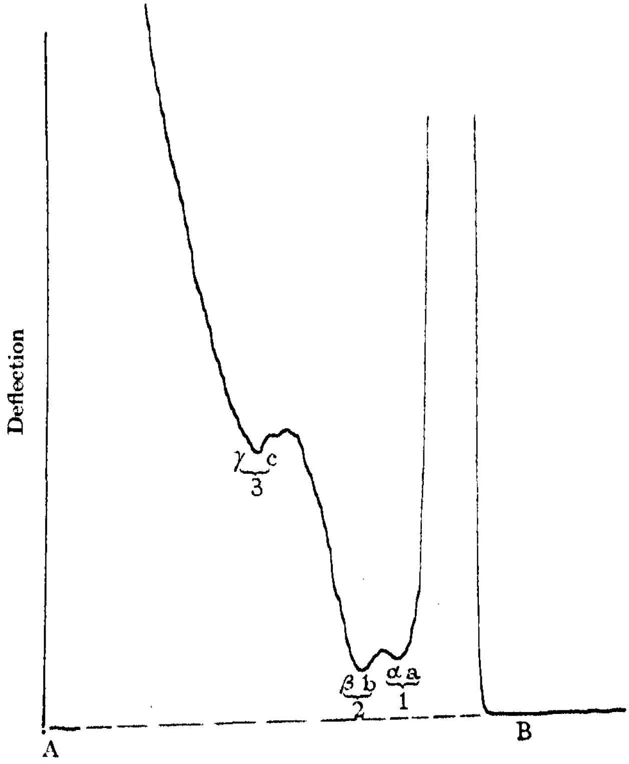


Fig. 21

The curve of fig. 21 has been used to determine the densities of the photographic images produced by the various components. The relative intensities can be calculated from the densities.

The derived intensity-time curve for the first four components at the point C of stroke 76 is shown in fig. 22, the intensity of the first component being arbitrarily taken as 100. It can be seen from this curve that the channel first has a luminous intensity of 100 and then fades to about 45 and increases again to about 115 for component 2. There now comes a sharp decrease to about 2, followed by two maxima of 18 and 2 for components 3 and 4 respectively. Component 5 has an intensity of 2, and component 6 is too faint for measurement. Components 4, 5 and 6 are not included in fig. 21.

It is evident that in the case discussed components 1 and 2 are of the same order of intensity and are the most important of the series. This conclusion appears to hold in general—the third and further components are much weaker than the first two.

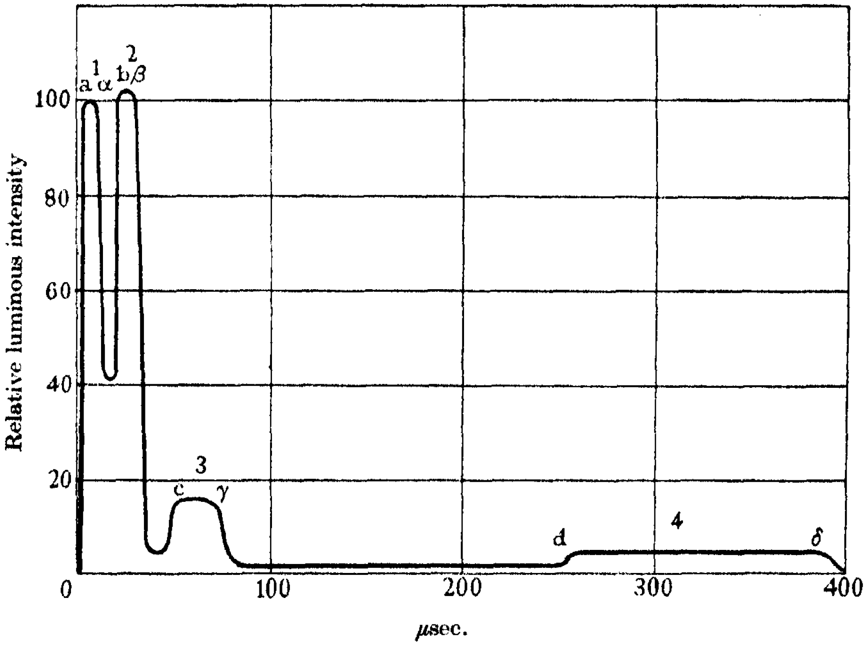


FIG. 22

This method involves the assumption that the spectrographic intensity distribution of the light remains constant till all the components are completed. It is also assumed that the channel width of the stroke remains constant, thus giving uniform exposure time for all the components at a fixed point on the stroke.

Special care has to be exercised that the photometric analysis is not made across bright streaks (Schonland and Collens 1934, p. 657) which are over-exposed.

A detailed study with improved spectrographic apparatus is in progress.

GENERAL DISCUSSION

This analysis of the luminosity-time relations in the return stroke shows clearly that within the limits of a resolving power of the order of 10 μsec., the only important variations in luminosity are those connected with the development of branches. Since it is natural to associate the luminosity in the channel with the current passing through it, we conclude that within the limits of our measurements important current variations in the channel arise solely from branches. Less important variations occur after the return stroke has reached the cloud. We can find no evidence of oscillations in the channel arising from the self-oscillation mechanism suggested by Simpson (1929). Such oscillations if they exist must have a frequency exceeding 100,000/sec. and appear to us to be unlikely in view of the mechanism of the return stroke.

The connexion between return stroke components and branches we have shown to be very close. Fig. 6 of § 5 indicates that a component of luminosity and hence of increased current arises shortly after a branch is reached. Our measurements are, however, such as to make it possible that this fresh component of current is called for after the branch has been carried some little way. For this increased current at branching points we can advance three causes:

(*a*) The additional call for current on account of the splitting of the channel along the branch and the main channel.

(*b*) The increased charge carried per unit length of a horizontal branch as compared with the vertical main channel.*

(*c*) The increasing freshness of the charge left by the leader as we move along a branch outwards.*

It is not proposed to discuss these causes in further detail here, as they are being dealt with in a more comprehensive survey of the discharge mechanism elsewhere.

We wish to thank the Lightning Research Committee of the South African Institute of Electrical Engineers under whose direction this research has been carried out. We are also greatly indebted to Professor B. F. J. Schonland, director of the Bernard Price Institute of Geophysical Research, for his assistance and advice in the preparation of this paper and to the Bernard Price Institute and Professor Paine of the Witwatersrand University for the hospitality extended to one of us (D. J. M.) as a guest researcher.

---

* These suggestions arose from discussions with Professor Schonland.

References

Boys 1929 *Nature, Lond.*, **124**, 54.
Lambert and Chalonge 1926 *Rev. Opt. (théor. instrum.)*, **5**, 404.
McEachron and McMorris 1936 *Gen. Elect. Rev.* **39**, 493.
Malan 1935 *S. Afr. J. Sci.* **32**, 129.
Malan, Schonland and Collens 1935 *Nature, Lond.*, **136**, 831.
Schonland and Collens 1934 *Proc. Roy. Soc.* A, **143**, 654.
Schonland, Malan and Collens 1935 *Proc. Roy. Soc.* A, **152**, 595.
Simpson 1929 *Journ. Inst. El. Eng.* **67**, 1275.
Walter 1935 *Phil. Mag.* **20**, 1144.

Rotatory Dispersion in the Amine Series
I—*d*-*Spiro*heptane-Diamine and its Derivatives

By the late T. M. Lowry, F.R.S. and W. C. G. Baldwin, M.A., Ph.D.

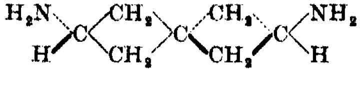
Laboratory of Physical Chemistry, Cambridge University,
and Physikalisch-Chemisches Institut, Heidelberg

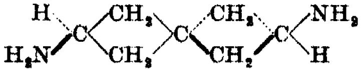
(*Communicated by R. G. W. Norrish, F.R.S.—Received* 21 *May* 1937)

I—Calculation of Rotatory Power

It has been shown by Janson and Pope (1936) that the diamines I and II are relatively easily resolved by means of the camphor-β-sulphonic acids. The substance II presents one of the simplest possible dissymmetrical structures, and quite apart from its chemical interest, is also of importance owing to its use by Max Born (1935) as an illustration of the possibility of quantitative predictions of rotatory power. The calculated specific rotatory power of the compound II was shown to agree with the observed value for the salt in water, but in Born's theoretical formula the rotatory power is dependent on the *8th power* of a distance, and for this reason the formula may find a wider application for localizing the vibrators in a molecule by use of the observed rotations.

In the work now described the preparations of Janson and Pope have been repeated and extended in order to provide: (*a*) Rotatory dispersion data from which accurate values for the active frequencies could be deduced, in place of the very approximate ones employed by Born. (*b*) The corresponding absorption spectra. (*c*) As many optically active examples of this interesting type as possible, with a view to further application to chemistry of Born's mathematical theory. For these purposes the following compounds were synthesized:

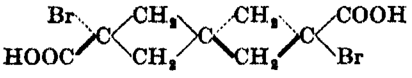
(I)  *l*-*Spiro*heptane-diamine,

(II)  *d*-*Spiro*heptane-diamine,

(cf. Lowry 1935, p. 391)

and the *d*-dihydrochloride, *d*-diacetyl-*spiro*base, *d*-dibenzoyl-*spiro*base, and *d*-diamidocarbonamino-*spiro*heptane. To these may be added two compounds prepared by Backer and co-workers (1931), viz. *d-spiro*heptane-dicarboxylic acid, and *d*-dibromo-*spiro*heptane-dicarboxylic acid (III):

(III)



(*d*-Dibromo*spiro*heptane-dicarboxylic acid)

In order to estimate the theoretical rotatory powers, suitable values are substituted in Born's simplified equation,

$$\chi = \frac{3{\cdot}49\times10^{-11}\,.\,\rho\,.\,(n^2+2)}{M\,.\,\lambda^2}\,.\left\{\frac{\lambda_1^3\,.\,\lambda_{11}^3}{(\lambda_{11}^2-\lambda_1^2)}\right\}^2.\,\frac{1}{L^8}\,.\,\frac{\frac{d}{L}\left(\frac{d^2}{L^2}-\frac{1}{2}\right)^2}{\left(\frac{d^2}{L^2}+\frac{1}{2}\right)^6},$$

where χ is the rotation in degrees per dm., ρ the density of the medium, n its refractive index, M the molecular weight, L the distance between the two vibrators at one end of the molecule, and d the effective length of the molecule. λ_1 and λ_{11} are characteristic wave-lengths associated with the end groups. In general it is best to take for the $>$CH wave-length a value of 1200 A, and for λ_{11} the value determined from the observed rotatory dispersion by means of a Drude equation. The values of L and d were deduced from a scale drawing in which the values C—N = 1·5, N—H = 1·0, C—H = 1·1, C—C = 1·5 A for the internuclear distances were used. For the amine, $L = 1{\cdot}87$ A and $d = 6{\cdot}15$ A, when the vibrators are localized at the centres of the N and H atoms. Assuming $n = 1{\cdot}5$, $\rho = 1{\cdot}1$, $\lambda_1 = 1200$ A, and putting $\lambda = 1360$ A, the value $[\alpha]_{5461} = 4{\cdot}3°$ is obtained, as compared with the observed value 7·8°. This comparison may indicate the necessity for slightly reducing L and d, thus localizing the C—H and C—NH_2 vibrators somewhere nearer to the rings, and to each other. For the dihydrochloride, for which $\lambda_{11} = 1934$ A, the rotation is still more reduced.

In dealing with substituted derivatives of the amine, the chief difficulty lies in localizing the second vibrator, since the attached groups are now large and anisotropic; for whilst one is probably justified in treating the groups CH_3 and NH_2, F and OH, NH_2 and H as approximately spherical, radicals such as CH_3CO—, C_6H_5CO—, etc., can clearly not be so classified. For this reason the theory cannot strictly be applied to such compounds, but instructive results are nevertheless obtained. The second vibrator has been considered as being at the mean position of the N and O atoms, which allows for free rotation about the C—N and N—C=O single bonds. For the acetyl base, L is then 2·2 A and $d = 6{\cdot}3$ A. Putting $\lambda_1 = 1200$ A and $\lambda_{11} = 1600$ A, and the other values as for the free base, $[\alpha]_{5461} = 0{\cdot}7°$, as compared with the observed value 3°. In order to obtain a better agreement, smaller values for λ_{11} and/or for *d* and *L* are necessary. Similar discrepancies are observed for the benzoyl compound. On substitution of a larger value for *M*, and using $\lambda_{11} = 1972$ A as derived from the corresponding Drude equation, $[\alpha]_{5461}$ is reduced to about 0·5°. The observed values show the benzoyl derivative to be more highly active than the acetyl, and the theory is thus directly opposed to experiment. This may find its explanation in the anisotropy of the groups, or in the introduction of the phenyl radicals, since with $\lambda_{11} = 1972$ A no fixed configuration gives sufficiently small values for *L* and *d*. Similarly, by localizing the second vibrator in the N or O atom only, *L* still remains too large both for the benzoyl derivative and for the carbonamide. Nevertheless, the dibromo acid gave a calculated value in excellent agreement with the value observed by Professor Backer, which was communicated to the authors after the calculation had been made—(C—Br = 2 A, $L = 2{\cdot}5$ A, $d = 6{\cdot}7$ A, $n = 1{\cdot}5$, $M = 342$, $\lambda_1 = 1800$ A, $\lambda_{11} = 2200$ A; $[\alpha]_D$ calc. = 7·45°, $[\alpha]_{5461}$ obs. = 7·8° for the sodium salt in water). In some cases the formula can therefore produce a value of the correct order, e.g. for the amine, the salt, the dicarboxylic acid, and for the dibromo acid, but in more complicated molecules the model is no longer applicable, and some other factor seems necessary to bring the calculated values within range of the observed.

Professor M. Born, having examined the present results, expressed his pleasure at the agreement obtained. He added, however, that before his assumption of isotropic oscillators was definitely accepted, he would like similar calculations to be made on the basis of anisotropy of the substituent groups. This can be done in a simple way by use of the molecular model proposed by Kuhn and Bein (1933), in which the chromophoric groups are regarded as possessing complete (linear) anisotropy. The relevant data for the base ($d = 6$ A, $a = 1$ A, $f = 1$, $\nu \equiv 5461$ A, $\nu_0 \equiv 1360$ A) were therefore substituted in Kuhn's equations (see Lowry 1935, p. 392). The specific rotation thus calculated, $[\alpha]_{5461} = 0{\cdot}5°$, is considerably lower than the observed value $[\alpha]_{5461} = 7{\cdot}8°$, although still of the correct order of magnitude, and the results show that the Born formula gives better agreement with experiment. We conclude, therefore, that the assumption of isotropic groups is justified.

II—ROTATORY DISPERSION OF *SPIRO*HEPTANE-DIAMINE DERIVATIVES

When the active absorbing region of a compound lies in the near ultraviolet, a detailed analysis of the optical properties can be made, owing to the possibility of measurements of circular dichroism, and therefore of the calculation of the rotatory dispersion on the basis of the equation of Lowry and Hudson (1933). Thus, the rotatory dispersions of some azides have been investigated by Kuhn and Braun (1930), of the aldehydes and ketones by Lowry and co-workers, and of chromophoric sulphur groupings by Levene (Levene and Rothen 1934), by Tschugaeff (1909), and by Lowry and Hudson (1933). Although large numbers of active amines have been prepared, detailed data for their rotatory dispersions are almost entirely confined to measurements on nicotine by Lowry and Lloyd (1929) and Lowry and Gore (1931) which were necessarily limited by the intense absorption of this compound; and on *spiro*heptane-diamine dihydrochloride for wave-lengths in the visual region only by Janson and Pope (1936). The inaccessibility of the absorption bands of simple amines presents a serious experimental limitation, but it has now been found possible to extend the range of observations for the *spiro* series to short ultra-violet wave-lengths.

(*a*) *Rotatory Dispersion of the Base*

The comprehensive data set out in Table V show that the dispersion of Pope and Janson's base is perfectly simple between 6708 and 2700 A. Although the aqueous solution absorbs strongly at 2200–2300 A, the dispersion equation $[\alpha] = 2{\cdot}1670/(\lambda^2 - 0{\cdot}01850)$ leads to the low value $\lambda_0 = 1360$ A for the wave-length of the active centre, which indicates that the longer wave-length absorption bands of this amine are quite inactive. This result is of great significance, and a full discussion is given in the succeeding paper of this series, where it is shown that, in general, bases undergo large changes of rotatory power on neutralization. When this occurs, the absorption bands of the amines at about 2300 A which can be attributed to the unbound electrons of the nitrogen atom become circularly dichroic. We have therefore concluded that *the rotatory power is due almost entirely to the absorption bands associated with the lone pair of electrons.* When the lone pair is used for bonding, as in the salts, these nearer bands disappear simultaneously with the disappearance of most of the rotatory power. The *spiro* base thus appears to behave as an exception in this respect, and further conclusive evidence of this is put forward in the subsequent paper, where the corresponding observations in the butylamine series are described.

(*b*) *Rotatory Dispersion of the Dihydrochloride*

Over the range 6708–2537 A, the dispersion of the hydrochloride of the base is also simple. The results can be represented with high accuracy by the equation $[\alpha] = 2{\cdot}261/(\lambda^2 - 0{\cdot}0374)$, $\lambda_0 = 1934$ A, whereas the equation

Table I—Rotatory Dispersion of *d*-Diamidocarbonamino-*spiro*heptane

$t = 20°$ C., $l = 6$ dm., $c = 1·467$ g. in 49·9 c.c. acetic acid

	λ	α	[α]	Dispersion ratios	
a	5780	0·496	2·814	*ab*	1·161
b	5461	0·577	3·270	*ac*	1·901
c	4358	0·945	5·350	*bc*	1·636

In water, $c = 0·2447$ g. per 50 c.c., $[\alpha]_{5461} = 1·70°$

Table II—Rotatory Dispersion of Diacetyl-*spiro*base in Alcohol at 20° C.

$c = 3·962$ g. per 100 c.c., $l = 6$ dm.

λ	α	[α]
5780	0·674	2·83 ± 0·08
5461	0·730	3·06 ± 0·10
4358	0·45	1·9
	(Photographic)	
4900	0·73	3·07
4736	0·68	2·86
4575	0·58	2·44
4552	0·53	2·23
4404	0·48	2·02
4255	0·38	1·60
4168	0·28	1·18
3932	0·00	0·00

Table III—Rotatory Dispersion of Dibenzoyl-*spiro*base in Pyridine at 20° C.

$c = 1·188$ g. per 100 c.c., $l = 6$ dm.

λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs}} - [\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$
6708	8·00	−0·43	6438	9·11	−0·12
6362	9·26	−0·20	6104	10·24	−0·15
5893	11·21	0·00	5780	11·64	−0·10
5700	12·34	0·24	5536	13·04	0·09
5461	13·32	−0·05	5219	14·71	−0·12
5209	14·86	−0·04	5153	15·29	0·00
5106	15·70	0·06	5086	15·99	0·23
4934	17·10	0·18	4811	17·95	−0·05
4722	18·94	0·14	4680	19·21	0·00
4602	19·80	−0·23	4554	20·35	−0·23
4358	22·01	−0·89	4325	22·88	−0·50
4143	24·85	−1·25	4045	26·23	−1·57
3940	27·65	−2·15	3821	29·08	−3·22

TABLE IV—ROTATORY DISPERSION OF *SPIRO*HEPTANE-DIAMINE HYDROCHLORIDE

Solvent = water at 20° C. c = 10·940 g. per 100 c.c.

λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}}-[\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}}-[\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}}-[\alpha]_{\text{calc.}}$
6708	5·30	−0·17	6497	5·75	−0·12	6438	5·94	−0·06
6362	6·07	−0·08	6104	6·67	−0·08	5893	7·35	0·05
5782	7·68	0·07	5780	7·68	0·07	5700	7·77	−0·09
5536	8·22	−0·18	5461	8·59	−0·07	5219	9·41	−0·20
5209	9·46	−0·19	5153	9·95	0·05	5106	9·92	−0·19
5086	10·10	−0·11	4934	10·91	−0·04	4811	11·65	0·00
4722	12·34	0·16	4680	12·70	0·26	4602	13·00	0·05
4554	13·49	0·20	4358	15·09	0·29		(Photographic)	
4442	14·39	0·28	4432	14·18	−0·03	4384	14·61	0·00
4350	15·09	0·20	4298	15·30	−0·03	4282	15·76	0·27
4263	15·76	0·11	4210	16·45	0·27	4152	16·69	−0·02
4107	17·37	0·17	4080	17·60	0·10	4007	18·50	0·18
3941	19·41	0·21	3833	20·56	−0·09	3769	21·70	0·10
3708	22·60	0·00	3630	24·00	0·05	3605	24·8	0·4
3522	25·81	−0·22	3441	27·9	0·0	3392	29·2	0·1
3329	30·1	−0·7	3284	31·6	−0·5	3232	32·8	−0·9
3205	34·1	−0·5	3178	35·6	0·1	3129	37·4	0·0
3108	39·2	1·0	3040	42·2	1·1	2967	44·1	−0·5
2900	49·5	1·1	2865	53·2	2·7	2788	56·9	0·8
2725	60·5	−0·8	2718	62·4	0·5	2686	66·5	1·3
2631	72·5	1·4	2582	78·9	1·8	2537	83·4	−0·3

TABLE V—ROTATORY DISPERSION OF *SPIRO*HEPTANE-DIAMINE IN WATER AT 20° C.

c = 7·845 g. per 100 c.c.

Solution prepared by adding the necessary quantity of caustic soda to the hydrochloride.

λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}}-[\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}}-[\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}}-[\alpha]_{\text{calc.}}$
6708	4·97	−0·04	6497	5·43	0·05	6362	5·74	0·13
6104	6·19	0·07	5893	6·57	−0·02	5782	6·76	−0·10
5780	6·76	−0·11	5700	7·14	0·07	5536	7·65	0·13
5461	7·78	0·03	5219	8·61	0·08	5153	8·80	0·03
5106	8·93	−0·02	4934	9·44	−0·19	4811	10·14	−0·03
4722	10·59	0·00	4680	10·96	0·15	4603	11·29	0·08
4554	11·41	−0·07	4358	12·50	−0·13		(Photographic)	
4575	11·35	0·00	4448	11·99	−0·09	4678	10·71	−0·09
4326	12·89	0·04	4216	13·51	−0·09	4412	12·25	−0·05
4150	14·10	0·00	4053	14·79	−0·09	4128	14·14	−0·12
3795	17·34	0·06	3792	16·65	−0·65	3906	16·05	−0·10
3470	21·00	−0·29	3365	23·05	0·17	3600	19·20	−0·30
3222	26·10	0·70	3140	26·90	−0·15	3240	25·61	0·57
3004	31·22	1·02	2900	33·30	0·30	3119	28·40	0·88
2767	37·45	0·15	2700	39·18	−0·62	2790	35·86	−0·69

given by Pope and Janson for the visual region only led to the value $\lambda_0 = 1452$ A. The important point to notice is that according to the Drude analyses, neither substance gives rise to circular dichroism in the near ultra-violet, although the base exhibits quite strong absorption in that region ($\epsilon = 7{\cdot}6$ at 2500 A, $\epsilon = 113{\cdot}3$ at 2229 A). For the hydrochloride the absorbing power is considerably reduced in consequence of the utilization for bonding of the electrons concerned ($\epsilon = 7{\cdot}52$ at 2240 A).

A comparison between the data for the base and the salt shows that neutralization of the base is not attended by any serious change in the rotatory dispersion—the dispersion centres are in the same spectral region, and the values of $[\alpha]^{20}_{5461}$ are comparable, viz. 7·78° and 8·59° respectively. For this reason, the absorption bands associated with the lone pair appear to have no connexion with the rotatory power. This is in direct contrast with many bases of Pasteur-type asymmetry, e.g. *l*-butylamine, ($[\alpha]^{20}_{5461} = -8{\cdot}63°$ for the pure base, $[\alpha]^{20}_{5461} = +1{\cdot}15°$ for the hydrochloride 23 % in water), where 80 % of the rotation is directly caused by the unbound nitrogen electrons, as will be shown in Part II. It seems reasonable to conclude that simple optically active bases may be divided into two types, namely: (i) the majority which show circular dichroism in the near ultra-violet bands of the nitrogen "lone pair", and (ii) a small number for which these bands are inactive, i.e. which behave like the *spiro* base. From the scanty data available in the literature, it appears probable that the latter class consists only of the following compounds: (i) The *menthylamines* investigated by Read (1930). (ii) *Tetrahydroquinaldine* (Read and Pope 1910), for which $\lambda_0 \div 1410$ A and $[\alpha]_D = 60{\cdot}04°$, whereas for the hydrochloride $\lambda_0 \div 1750$ A, and $[\alpha]_D = 66{\cdot}47°$. (iii) *d-αβ-Diamino-γ-brompropane* for which $[\alpha]_D = 5{\cdot}20°$ (base), and $[\alpha]_D = 7{\cdot}27°$ (hydrochloride). (iv) *d-Aminobenzyl-β-naphthol*, $[\alpha]_D$ (base) $= 58{\cdot}84°$, $[\alpha]_D$ (hydrochloride) $= 52{\cdot}89°$. (v) *d-αβ-Diamino-γ-methoxypropane*, $[\alpha]_D = 8{\cdot}19°$ (base). $[\alpha]_D = 9{\cdot}04°$ (salt). (vi) A series of amines described by Levene and Marker (1936). Most amino acids also seem to behave similarly, but as these compounds exhibit internal salt formation, and consequently do not possess typically basic properties, they are best omitted from the present discussion.

(c) *Rotatory Dispersion of the Diacetylbase*

This substance in alcoholic solution exhibits anomalous rotatory dispersion, as shown in fig. 1, since two absorption bands are now active, and give rise to components of the circular dichroism which are of opposite sign. The band of shorter wave-length has slightly the larger activity, although the two partial rotations approximately cancel one another, so that only a small residual rotation is observed. This curve is of interest because of its great resemblance to those for tartaric ester, as recorded by Lowry and Cutter (1922), and for 50 % tartaric acid in water (Lowry and Austin 1922). On this basis the two bands may be attributed respectively to the N—H radical (at approximately 1700 A) and to the acetyl groups (at approximately 2200 A), so that there is no need to assume any effect from a band at the latter wave-length due to lone pairs on the two nitrogen atoms.

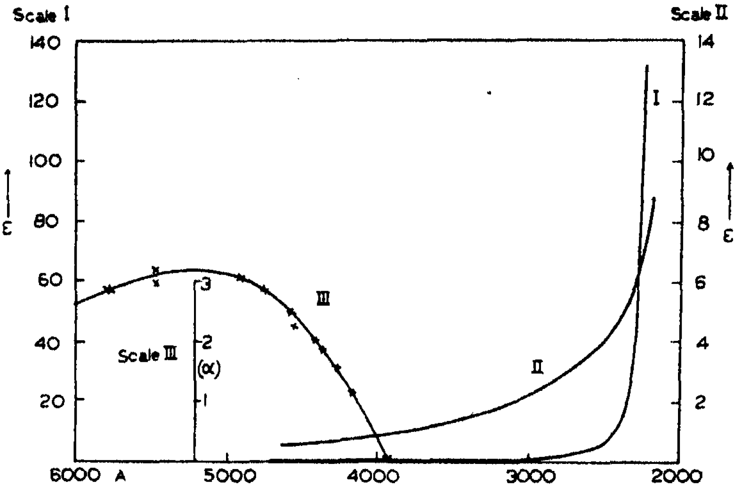


FIG. 1—Curve I: absorption spectrum of aqueous *spiro*heptanediamine on scale I; $c = 0{\cdot}949$ M. Curve II: absorption spectrum of the corresponding dihydrochloride on scale II; $c = 0{\cdot}213$ M. Curve III: rotatory dispersion of diacetyl-*spiro*heptanediamine in alcohol, on scale III.

(*d*) *Rotatory Dispersion of the Dibenzoyl Base*

Measurements for this compound are limited in extent and accuracy by its low solubility even in such an unsuitable solvent as pyridine. The rotatory dispersion can, however, be represented by the equation

$$[\alpha] = 3{\cdot}463/(\lambda^2 - 0{\cdot}0389),$$

which, although not perfect, demonstrates clearly that it arises in a strong absorption band with $\epsilon = 18000$ at 2350 A. This band undoubtedly contains at least two absorption systems associated respectively with the N—H and benzoyl groups, and the optical activity probably has its origin in the benzoyl bands. It should be noticed that the "step-out" at 2800 A due to the phenyl rings, which is shown in the absorption curve for this compound in fig. 2, is definitely inactive. This appears to be in contrast to the behaviour of mandelic acid, for which a strong induced dissymmetry was attributed by Kuhn (1930) to this source. In the acid, however, the phenyl group is attached to an asymmetric carbon atom, whereas in the *spiro* compound it forms part of the configuration. The two cases are therefore not strictly comparable. In this connexion it is important to remark that all seven members of this *d*-spirane series so far prepared are uniformly dextrorotatory, in contrast to the butylamine series where reversals are common. The reason for this uniformity is that the influence of a fixed dextro-rotatory nucleus is always impressed in the same way, but to different extents, on the absorption bands of the attached groups.

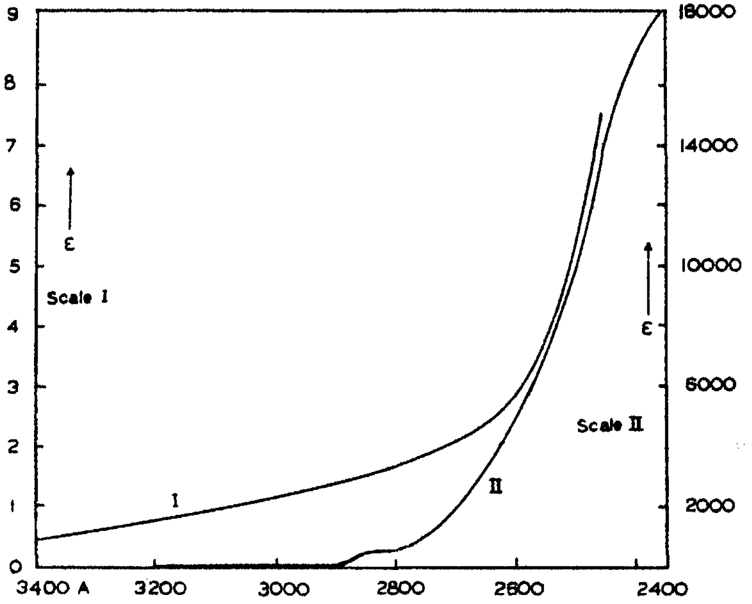


Fig. 2—Curve I: absorption spectrum of diacetyl-*spiro*heptanediamine in alcohol on scale I; c = 0·1375 M. Curve II: absorption spectrum of dibenzoyl-*spiro*heptanediamine in methyl alcohol, on scale II; c = 0·015–0·000025 M.

(*e*) *Rotatory Dispersion of d-Diamidocarbonamino*spiro*heptane*

A solution of this compound in water gave the low value $[\alpha]_{5461} = 1{\cdot}7^\circ$. The rotation was higher for acetic acid solutions, $[\alpha]_{5461} = 3{\cdot}27^\circ$. The dispersion in each case was centred at 2050 A, and can therefore be attributed to the —CO— groups.

(*f*) *Absorption Spectra*

The large decrease in absorption on neutralization of a base with hydrochloric acid has already been discussed. The data for the *spiro* base and its salt are shown in fig. 1 and those for the acyl derivatives in fig. 2.

EXPERIMENTAL

The general methods of preparation used by Janson and Pope were followed, but a few minor modifications were introduced. The *spiro*-dicarboxylic acid was found to crystallize free from troublesome tar when 50% glacial acetic acid and 50% strong HCl was used as the solvent. Eventually some 15 g. of pure *d* salt was obtained, and 7 g. of the laevo product.

*d-Dibenzoyl*spiro*base*. The product of benzoylation was twice crystallized from alcohol/benzene. Its melting-point, 242° C., was identical with the value given by Janson and Pope, who recorded $[\alpha]_{5461} = 14{\cdot}5°$ for their specimen in pyridine.

*d-Diacetyl*spiro*base*. The free base was refluxed with a slight excess of acetic anhydride, and then evaporated to dryness. The solid was crystallized three times from purified pyridine, and once from benzene/alcohol. Its melting-point was 215–216° C., i.e. 20° C. higher than that of the *dl* compound made by Janson and Pope.

Pure d-dihydrochloride. The impure material was dissolved in wet alcohol, filtered, and evaporated with pure benzene until it crystallized. This procedure was repeated.

*d-Diamidocarbonamino*spiro*heptane*. The solid was crystallized twice from water, once from methyl alcohol, and once from aqueous acetic acid. Melting-point 260° C.

Pure l-spiro*heptane-diamine*. The wet base was heated for some time with its own weight of potash in small pieces, and distilled *in vacuo*. The distillate was diluted with chloroform and dried over anhydrous sodium sulphate. Two more distillations yielded an oil, with a tendency to fume, which solidified in ice/salt, and on slow warming melted at 17·5° C. Janson and Pope recorded 7–8° C. as the melting-point of the *dl* base.

1 : 1′-*Dihydroxy*spiro*heptane*. 3·5 g. of pure *d*-dihydrochloride was dissolved in water in a 50 c.c. flask, a little HCl being added. Double the calculated amount of sodium nitrite solution was slowly run in. When the nitrite had decomposed, the solution gave $[\alpha]_{5461} = 6{\cdot}8°$. Vacuum distillation from acid gave no product, but alkaline distillation gave unchanged base in good yield.

The authors wish to thank Professor Sir William Pope for his interest in the work, Dr. F. G. Mann for the gift of 1 kg. of pentaerythritol, the Chemical Society and the Goldsmiths' Company for research grants.

SUMMARY

1—A series of *spiro*heptane derivatives has been prepared, and their theoretical rotatory powers calculated according to Born's theory. Values of the correct order are only obtained when no long side-chain is present.

2—The absorption spectra of *spiro*heptane-diamine, its salt, and acetyl and benzoyl derivatives are recorded, and their significance discussed.

3—The rotatory dispersions of these compounds and of diamidocarbonamino*spiro*heptane are analysed, and correlated with the absorption measurements.

4—The large changes of rotatory power of most amino bases, which are observed on neutralization, are caused by the saturation in this process of the "lone pair" of electrons on the nitrogen atom, which usually give rise to circularly dichroic absorption bands.

5—Amino bases are divided into two classes according as the absorption bands in the 2300 A region are active or inactive. When these bands are inactive, little change in rotatory power takes place on neutralization of the base, but such behaviour is exceptional. The *spiro*heptane base falls in the latter class.

REFERENCES

Backer and Schurink 1931 *Rec. Trav. chim. Pays-Bas*, **50**, 921.
Baldwin, Wolfrom and Lowry 1935 *J. Chem. Soc.* p. 696.
Born 1935 *Proc. Roy. Soc.* A, **150**, 104.
Hudson, Wolfrom and Lowry 1933 *J. Chem. Soc.* p. 1179.
Janson and Pope 1936 *Proc. Roy. Soc.* A, **154**, 53.
Kuhn 1930 *Trans. Faraday Soc.* p. 307.
Kuhn and Bein 1933 *Z. phys. Chem.* B, **22**, 406.
Kuhn and Braun 1930 *Z. phys. Chem.* B, **8**, 281.
Levene and Marker 1936 *J. Biol. Chem.* **115**, 267.
Levene and Rothen 1934 *J. Chem. Phys.* p. 681.
Lowry 1935 "Optical Rotatory Power", p. 385. Longmans.
Lowry and Austin 1922 *Philos. Trans.* A, **222**, 249.
Lowry and Cutter 1922 *J. Chem. Soc.* p. 532.
Lowry and Gore 1931 *J. Chem. Soc.* p. 319.
Lowry and Hudson 1933 *Philos. Trans.* A, **232**, 117.
Lowry and Lishmund 1935 *J. Chem. Soc.* p. 709.
Lowry and Lloyd 1929 *J. Chem. Soc.* p. 1771.
Lowry and co-workers 1933 et seq. *J. Chem. Soc.* p. 1185.
Read 1930 *Trans. Faraday Soc.* p. 441.
Read and Pope 1910 *J. Chem. Soc.* p. 2199.
Tschugaeff 1909 *Ber. dtsch. chem. Ges.* **42**, 2244.
— 1910 *Z. phys. Chem.* **74**, 503.

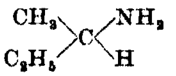
Rotatory Dispersion in the Amine Series
II—*l*-Secondary Butylamine and its Derivatives

By W. C. G. Baldwin, M.A., Ph.D.

(*Communicated by R. G. W. Norrish, F.R.S.—Received* 21 *May* 1937)

In the preceding paper, nitrogenous bases were divided into two classes according as the first amino-absorption bands, which they exhibit at about 2200 A, do or do not contribute to the optical activity of the molecules. Thus, for *spiro*heptane-diamine, the absorption spectrum showed a rapid rise in intensity caused by the presence in the molecule of two lone-pairs of electrons associated with the nitrogen atoms. On neutralization of the base, this strong absorption at 2300 A disappeared, but the rotatory dispersion remained almost unchanged. For both base and salt the *active* absorption region was between 1950 and 1400 A, and the bands of the amine at longer wave-lengths were not circularly dichroic. A test of the conditions prevailing in any particular case was specified on the basis of the behaviour of the specific rotation on neutralization, since, when a large change of magnitude or sign occurs, the bands of longer wave-length must play an important role. Care is necessary in the use of this rule, because solvents such as water may co-ordinate with the amine, causing the behaviour of the free base to resemble that of its salts. When, however, a non-polar solvent is used for the base, and the change is observed, it is practically certain that this interpretation is correct.

In the present paper, a detailed account is given of the corresponding phenomenon as shown by a base whose optical properties are more typical of this class of compounds, namely, secondary butylamine:



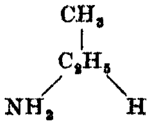
$$\begin{matrix} CH_3 \\ C_2H_5 \end{matrix} \!>\! C \!<\! \begin{matrix} NH_2 \\ H \end{matrix}$$

The effects of neutralization have already been discussed by Rule (1930), but his attention was confined to the relationship between polarity and rotatory power, and the question of circularly dichroic absorption bands was ignored. Since, however, polarity is usually manifested in the groups with absorption bands nearest to the visual region, the two arguments run to some extent parallel to one another, but the present point of view is more fundamental.

Ignoring for the present the amino-acids, the simple bases showing no "acid effect" can be divided into three groups: (*a*) Menthylamines, tetrahydroquinaldine, and *spiro*heptane-diamine. (*b*) $\alpha\beta$-diamino-γ-bromopropane, aminobenzyl-β-naphthol, and $\alpha\beta$-diamino-γ-methoxypropane. (*c*) A series described by Levene and Marker (1936) of general formula $NH_2.(CH_2)_m.CHMe.(CH_2)_nMe$. Consideration of the three substances under (*a*) indicates the tendency of "ring dissymmetry" to cause no near circular dichroism in bases, i.e. no "acid effect". Similarly, group (*c*) apparently shows the inability of the asymmetric grouping to function as an activator of the lone-pair over great distances. The substances under (*b*) are, for no apparent reason, some of the few exceptions to the rule among the simple amines. With a view to elucidating this fact, which cannot be explained, it is proposed to make at a later stage a full investigation of propylene-diamine and its derivatives.

The substance chosen for the present work, *l*-secondary butylamine, has one of the simplest possible structures producing optical activity, since the amino group is directly attached to the asymmetric carbon atom, and is thus fully exposed to the electrical anisotropy of the molecule. It is not surprising, therefore, to find that the "acid effect" is very marked.

The laevo-base, prepared and resolved by the method of Pope and Gibson (1912), was found to have $[\alpha]_D^{20} = -7{\cdot}45°$, and $d_4^{20} = 0{\cdot}7205$, whereas the laevo-hydrochloride gave $[\alpha]_D^{20} = +0{\cdot}96°$. This striking change of sign may be interpreted in two ways, one of which, the "electronic" interpretation, has been mentioned above. Another, but tentative point of view, is provided by Boys's definition of the configuration of the amyl alcohols (1934), on the basis of which it would be concluded that *l*-butylamine has the configuration



where the ethyl group is above the plane of the paper ("molecular" interpretation). In this model the negative sign of the rotation is defined by the anticlockwise order of the groups as they diminish in size. When the amino group is ionized, the order of size, and sign of rotation are both reversed. At the same time the rotation of the salt is small owing to the electronic similarity between the $\overset{+}{N}H_3$ and the CH_3. On acylation of the amino radical, the latter must, as the largest in the molecule, take the place of the ethyl radical above the paper, and consequently, acylbutylamines should be, and are, uniformly laevo-rotatory. All substituted secondary butylamines follow Boys's rule, with one exception, namely

$$CH_3.CH(NH_2).CH(OC_2H_5)_2,$$

described by Fischer and Kametaka (1909). The rule also breaks down for propylene diamine, since it predicts that the rotatory power of the dihydrochloride should be opposite in sign to that of the base, whereas, in reality, it is not. Despite the success of the Boys model for many cases, it is therefore clear that the best method of approach to the problems offered by the bases is that which is founded on the consideration of active absorption bands. From this point of view, elimination of circular dichroism postulates a reduction in the magnitude, but not necessarily a reversal of the sign of the rotation, just as is observed for propylene diamine. The point can be readily appreciated from a study of Table I, compiled from the literature.

TABLE I

Base	$[\alpha]_D$	$[\alpha]_D$ HCl
l-Butylamine	− 7·45	0·96 (Present values)
Diethyl-*d*-butylamine	85·8	4·24 (Present values)
Nicotine	− 169·3	15·45
l-α-Phenylethylamine	− 39·9	− 5·3
Amylamine	− 5·86	0·0 (Marckwald)
		− 0·21 (Levene $[M]_D$)
l-Propylenediamine	− 28·04	− 4·04
N-Et-*l*-α-Phenylethylamine	− 61·2	− 12·2
N-Pr	− 69·1	− 24·3
N-Benzyl	− 40·1	− 9·1
d-α-*p*-Tolylethylamine	36·57	3·4 (Sulphate)
l-*p*-Et-Phenyl	− 31·7	− 6·3 (Sulphate $c = 3·34$)
		− 20·6 (Sulphate $c = 0·41$)
		− 30·8 (Oxalate, hydrolysis)
Pseudoephedrine	51·2	− 33·8
l-Dihydro-α-methylindole	7·22	1·68
l-Narcotine	− 199·3	+
Cocaine	− 15·75	− 57·5
l-Hyoscyamine	− 23	32
α-Aminocamphoroxime	60·5	36·7
Bornylamine	46·2	16·0 (HI)
Methylbornylamine	81·0	26·6 (HI)
d-2-Amino-octane	5·41 $[M]_D$	− 6·44 $[M]_D$

Absorption Spectra—In common with all simple nitrogenous bases, butylamine exhibits strong absorption at about 2300 A which disappears on neutralization, as shown in fig. 1. Other examples of this behaviour are found in ethylene-diamine, primary, secondary, and tertiary alkylamines, piperidine, and ammonia (Ley and Arends 1932; Bielecki and Henri 1913;

Ley and Ulrich 1909; Hartley and Dobbie 1900). Of these, ammonia is an excellent example, since the reduction in intensity of absorption, which occurs on solution in water, is due to interaction with the solvent. Solutions in hexane gave the typical amino absorption, which is now attributed to the lone-pair (Ley and Arends 1932). In general, amino acids lack this property on account of internal co-ordination, and must be treated separately.

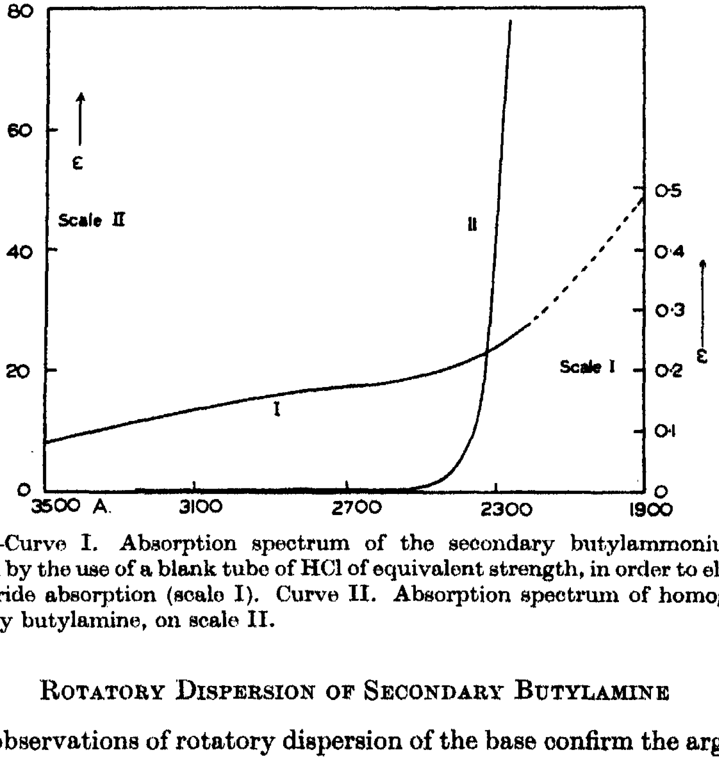


Fig. 1—Curve I. Absorption spectrum of the secondary butylammonium ion, obtained by the use of a blank tube of HCl of equivalent strength, in order to eliminate the chloride absorption (scale I). Curve II. Absorption spectrum of homogeneous secondary butylamine, on scale II.

Rotatory Dispersion of Secondary Butylamine

The observations of rotatory dispersion of the base confirm the argument set out above. For the pure liquid, the dispersion between 6708 and 3485 A can be expressed accurately by a Drude equation with two terms, corresponding to active frequencies at 2300 and 1792 A. These values are in excellent agreement with Mulliken's observations (1935) of the two chief absorption regions of aliphatic amines, namely at 2400–2000 and at 1800–1500 A respectively. The two Drude terms have opposite signs, and the dispersion must be classed as anomalous, although the anomaly actually occurs just outside the range of observations, as is shown by the following calculated wave-lengths: Reversal of sign, 2580 A; maximum rotation at 2975 A; point of inflexion in dispersion curve, 3330 A. In order to obtain direct observations of the anomaly, experiments were made with an aqueous solution. The active wave-lengths were then at 2258 and 1761 A, both showing the slight displacement to be expected when a hydroxylic solvent is used. The wave-lengths of the anomaly for the aqueous solution are in good agreement with values calculated from the corresponding Drude equation: maximum 3185 A (calc.) 3150 A (obs.); inflexion at 3625 A (calc.) 3500–3600 A (obs.); reversal of sign at 2692 A. The difference between this type of dispersion and that of *spiro*heptane-diamine is very striking, not so much because of the anomaly, but because the first active absorption region for butylamine is 900 A nearer to the visual region.

ROTATORY DISPERSION OF BUTYLAMINE HYDROCHLORIDE

Measurements of the rotation for the hydrochloride in aqueous solution in the visible and near ultra-violet regions lead to a Drude equation in which the active wave-length is 2262 A; but the absorption at this wave-length due to the butylammonium ion is only $\epsilon = 0{\cdot}27$, and the ion shows no appreciable absorption at wave-lengths above 1900 A. It appears likely, therefore, that the simple Drude equation used does not completely represent the facts, and that the partial rotation expressed by it is a resultant of several higher frequency terms. It is highly improbable that circular dichroism of any magnitude could be manifested by such weak absorption, and the dispersion, therefore, should probably be considered as pseudo-simple. This conclusion is confirmed by the magnitude of the "acid effect" described above, which is the most important factor in explaining the behaviour of these compounds. The following paragraph gives further strong reasons supporting this argument.

ROTATORY DISPERSION OF N-DIETHYL-*d*-SECONDARY BUTYLAMINE

The dispersion of diethyl-butylamine is in striking contrast to that of the parent substance, since the anomaly is no longer present. The Drude equation leads to a value $\lambda_0 = 2000$ A at the higher frequency end of the system of bands produced by the lone pair, between 2300 and 2000 A. On neutralization of this base, the value of $[\alpha]_D$ falls from 85·8° to 4·24°, and the calculated dispersion centre moves from 2000 to 1500 A, owing to the more saturated state of the salt.

ROTATORY DISPERSION OF ACETYL- AND BENZOYL-BUTYLAMINES

The dispersion of benzoyl-butylamine strongly resembles that of the corresponding *spiro*heptane-diamine derivative. A one-term Drude equation expresses the results satisfactorily over the range 6708–3290 A, and the wave-length of the active centre derived from it is 2057 A, which is in good agreement with the value 1972 A obtained for the *spiro*-compound. Absorption in this region is attributed to the carbonyl part of the benzoyl group, and the results therefore indicate that most of the activity of these two compounds is caused by the induced dissymmetry of the benzoyl radicals. No term contributed by the nitrogen electrons is detectable. Such a result is not surprising when it is remembered that a base loses most of its typical properties on acylation.

In contrast to diacetyl-*spiro*heptane-diamine, acetyl-butylamine gives no sign of anomaly in its dispersion. Between 6708 and 3290 A the ideal Drude equation again has only one term, with $\lambda_0 = 1631$ A. This result is to be expected in view of the low values of molecular extinction coefficients at 2400 A for the acetyl-*spiro*base as compared with those of the benzoyl compound, since the circularly dichroic acetylamine bands would seem to fall in a region of lower wave-lengths; but the unique development of anomalous dispersion in diacetyl-*spiro*heptane-diamine is very curious.

ROTATORY DISPERSION OF SECONDARY BUTYL-UREA

The dispersion of this substance is completely analogous with that of the corresponding *spiro*-compound. As before, the rotatory power is caused by the induced dissymmetry of the urea residue, centred at a mean wave-length of 1845 A, as compared with 2050 A for the *spiro*-urea. This value serves as further demonstration that substitution of the amino group causes it to lose both its typical basic properties, and its dispersion characteristics. The nitrogen electrons no longer show any effect in the middle ultra-violet, and the substituting group now controls the rotatory power.

Nicotine—The extensive researches of Lowry into the optical properties of nicotine and its derivatives have established the following facts: (*a*) The rotatory power of the base is very high, and the dispersion is centred at 2525 A. (*b*) The salts have smaller rotatory powers, opposite in sign, with the active centre at 1500 A. (*c*) Both nicotine and its zinc chloride compound have an intense absorption band at 2650 A, which, in the case of the base, was identified as the dispersion centre. In the salts this band is apparently inactive, and the dispersion is controlled by the frequencies in the Schumann region only. (*d*) A similar "acid effect" occurs in the methiodide, which has two methyl groups on the nitrogen atom. The changes observed on formation of the methiodide are therefore *not* due to the development of a new asymmetric centre in the N^+, such as occurs in the formation of the nicotinium ion. (*e*) The pyridine nucleus is assumed to be the source of the chief active absorption bands in nicotine, but in the ion and methiodide, the same bands *although present* are inactive. (*f*) The iso-methiodide containing tervalent pyrrolidine nitrogen, conformed to the original nicotine type. Utilization of the lone-pair, either in the methiodide or dimethiodide, resulted in the characteristic change both of sign and dispersion. (*g*) Chlornicotine and methylnicotone conform to the type of the free base; both contain the pyrrolidine lone-pair. The dispersion of these two compounds is centred at 2550 and 2668 A respectively. Lowry suggested that the agreement between these figures indicated that the active absorbing group in each case is the same, i.e. the pyridine nucleus. The ketonic group is inactive, although it gives rise to an absorption band.

On the basis of the results obtained with butylamine and its derivatives, an alternative interpretation of these phenomena can now be proposed. Thus: (i) The pyridine and benzene absorption bands are of a type which rarely show circular dichroism. (ii) The rotatory dispersion of nicotine is controlled chiefly by a frequency connected with the pyrrolidine lone-pair, which gives rise to a dichroic band buried in the strong but inactive pyridine absorption. On ionization of the pyrrolidine nitrogen, the absorption by the pyridine nucleus remains, and even increases in intensity, but the dichroic portion vanishes. The Schumann frequencies then become the centre of the optical activity. (iii) Similarly, in the methiodides, *whenever the pyrrolidine lone-pair is unsaturated*, the frequency at 2500–2600 A appears in the equations of rotatory dispersion, and *when the nitrogen is ionized it vanishes*. (iv) Methylnicotone has a dispersion controlled by the same active centre, but the ketonic group is also slightly affected, and the resultant dispersion is pseudo-simple. The value of the active wave-length obtained from the Drude equation falls at 2668 A and represents a weighted mean of the two absorbing regions at 2900 and 2400 A respectively.

EXPERIMENTAL

Resolution of Secondary Butylamine

The bromocamphor-sulphonate of the *dl* base was fractionally crystallized twenty-seven times from absolute ethyl acetate, and the optical purity tested from time to time by distilling a sample with baryta, determining the rotatory power, and estimating the concentration by titration. Determinations of melting-point were valueless, presumably owing to the existence of two crystalline modifications. On one occasion a dry sample melted at 98° C., and, on standing, the melting-point rose to 157° C. After another crystallization it fell again to the lower figure.

TABLE II—ROTATORY DISPERSION OF SECONDARY BUTYL-UREA IN WATER AT 20° C.

c = 2·015 g./100 c.c.

	λ	(α)	Ratios
a	5780	−27·3	1·139 *ab*
b	5461	−31·08	1·684 *bc*
c	4358	−52·4	1·919 *ac*

TABLE III—ROTATORY DISPERSION OF HOMOGENEOUS *l*-SECONDARY BUTYLAMINE AT 20° C.

$$100\,[\alpha] = 163{\cdot}0/(\lambda^2 - 5{\cdot}290) - 409/(\lambda^2 - 3{\cdot}212).$$
$$\lambda_{\text{I}} = 2300\text{ A};\ \lambda_{\text{II}} = 1792\text{ A}.$$

λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$
6708	− 5·76	−0·07	6362	− 6·34	0·01
6104	− 6·90	0·00	5893	− 7·45	0·00
5782	− 7·65	0·08	5780	− 7·65	0·10
5700	− 7·97	0·00	5536	− 8·43	0·03
5461	− 8·63	0·08	5219	− 9·54	0·02
5153	− 9·76	0·09	5106	− 9·95	0·09
4934	−10·76	0·04	4811	−11·60	−0·23
4722	−11·95	−0·11	4680	−12·21	−0·13
4602	−12·57	−0·08	4554	−12·81	−0·04
4370	−13·89	0·07	4358	−13·99	0·01
4283	−14·53	−0·02	4195	−15·23	0·00
4096	−16·06	−0·08	3993	−16·89	−0·10
3901	−17·59	0·02	3849	−18·17	−0·06
3842	−18·28	−0·08	3804	−18·56	0·00
3713	−19·01	0·51	3660	−19·85	0·21
3556	−20·68	0·48	3485	−21·52	0·43

Acetyl-butylamine. This substance was not isolated. A standard flask of appropriate size was taken, and rather more than half-filled with very pure acetic anhydride. The pure amine was run into this weighed quantity of anhydride through a very fine capillary tube, the flask meanwhile being cooled in ice. A very violent reaction occurred, and the acetyl derivative was immediately formed. The quantities were adjusted so that there was a slight excess of unused anhydride.

Benzoyl-butylamine (Pope and Gibson 1912). This was purified by two crystallizations from optically pure cyclohexane. M.P. 92° C.

Secondary Butyl-urea (Urban 1904). The impure material was crystallized twice from water. M.P. 168–169° C. (166° C., Urban). Analysis: nitrogen 24·27 % obs., 24·2 % calc.

d-Diethyl-butylamine (Leithe 1930). The picrate was crystallized four times from alcohol, until its melting-point was constant at 117° C. The base was steam distilled out from the alkaline picrate solution, and dried by solid caustic soda. Its purity was estimated as better than 99·6 %. B.P. 130–131° C. $d_4^{20} = 0{\cdot}753$. $[\alpha]_D^{20} = 85{\cdot}8°$. Leithe found $[\alpha]_D^{15} = 89{\cdot}3°$.

The author expresses his indebtedness to the late Professor T. M. Lowry, F.R.S., and thanks Dr. F. B. Kipping for the loan of butylamine. He also gratefully acknowledges grants from the Goldsmiths' Company and from the Chemical Society.

TABLE IV—ROTATORY DISPERSION OF *l*-SECONDARY BUTYLAMINE IN WATER AT 20° C.

$c = 8{\cdot}89$ g./100 c.c. in 6 or 1 dm. tube.

$$100\,[\alpha] = 176{\cdot}1/(\lambda^2 - 5{\cdot}100) - 341{\cdot}8/(\lambda^2 - 3{\cdot}099).$$

λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$
6708	− 3·69	0·04	6497	− 4·01	−0·03
6438	− 4·07	−0·01	6362	− 4·26	−0·09
6104	− 4·52	0·01	5893	− 4·86	0·00
5782	− 5·08	−0·04	5780	− 5·04	0·00
5700	− 5·27	−0·07	5536	− 5·49	0·01
5461	− 5·66	0·01	5219	− 6·26	−0·06
5209	− 6·28	−0·07	5153	− 6·47	−0·13
5106	− 6·56	−0·08	5086	− 6·56	−0·06
4934	− 6·88	0·04	4912	− 6·92	0·04
4811	− 7·44	−0·19	4800	− 7·54	−0·24
4722	− 7·54	0·01	4680	− 7·80	−0·11
4678	− 7·87	−0·19	4602	− 7·85	0·08
4554	− 8·14	−0·02	4358	− 8·77	0·01
4304	− 9·03	0·00	4225	− 9·32	0·04
4143	− 9·69	−0·01	4142	− 9·67	0·01
4053	−10·07	0·00	3965	−10·53	−0·10
3930	−10·23	0·48	3900	−10·80	0·00
3871	−10·91	0·03	3815	−11·20	0·01
3700	−11·81	−0·03	3471	−12·94	−0·14
3290	−14·06	−0·63	3150	Maximum	

TABLE V—ROTATORY DISPERSION OF *l*- SECONDARY BUTYLAMINE HYDROCHLORIDE

$t = 20°$ C. $c = 22{\cdot}84$ g./100 c.c. water. $l = 6$ dm.
$[\alpha] = 0{\cdot}2923/(\lambda^2 - 0{\cdot}05116)$.

λ	$[\alpha]_{obs.}$	$[\alpha]_{obs.} -$ $[\alpha]_{calc.}$	λ	$[\alpha]_{obs.}$	$[\alpha]_{obs.} -$ $[\alpha]_{calc.}$
6708	0·77	0·04	6497	0·80	0·01
6438	0·79	−0·01	6362	0·83	0·00
6104	0·89	−0·02	5893	0·96	−0·03
5780	1·00	−0·03	5700	1·06	0·00
5536	1·11	−0·03	5461	1·15	−0·03
5219	1·31	−0·01	5209	1·31	−0·02
5153	1·36	0·00	5106	1·44	0·05
5086	1·45	0·04	4934	1·58	0·06
4811	1·62	0·00	4800	1·67	0·04
4722	1·73	0·03	4678	1·76	0·02
4602	1·86	0·04	4584	1·75	−0·08
4554	1·91	0·04	4477	1·90	−0·06
4433	1·97	−0·04	4358	2·19	0·08
4353	2·12	0·01	4253	2·26	0·01
4138	2·41	−0·02	4067	2·55	−0·01
4030	2·70	0·07	3951	2·85	0·06
3852	3·03	0·02	3792	3·25	0·09
3713	3·43	0·05	3668	3·61	0·11
3584	3·79	0·01	3535	3·98	0·02
3471	4·16	−0·05	3450	4·29	−0·02
3400	4·51	−0·03	3355	4·68	−0·09
3310	4·86	−0·14	3295	5·03	−0·07
3250	5·43	0·06	3134	6·13	−0·08
3100	6·57	0·06	3048	7·00	0·00
2999	7·18	−0·37	2975	7·61	−0·22

TABLE VI—ROTATORY DISPERSION OF ACETYL-*l*-SECONDARY BUTYLAMINE AT 20° C.

$c = 34{\cdot}40$ g./100 c.c. acetic acid.
$[\alpha] = -5{\cdot}567/(\lambda^2 - 0{\cdot}02660)$.

λ	$[\alpha]_{obs.}$	$[\alpha]_{obs.} -$ $[\alpha]_{calc.}$	λ	$[\alpha]_{obs.}$	$[\alpha]_{obs.} -$ $[\alpha]_{calc.}$
6708	−13·21	−0·05	6497	−14·10	−0·01
6362	−14·72	0·00	6104	−16·08	0·00
5893	−17·30	0·07	5782	−18·10	0·00
5780	−18·17	−0·05	5700	−18·66	0·00
5536	−19·97	−0·16	5461	−20·50	0·00
5219	−22·66	0·01	5153	−23·26	0·04
5106	−23·89	−0·10	4934	−25·70	0·00
4912	−26·00	−0·08	4811	−27·52	−0·32
4722	−28·51	−0·19	4720	−28·62	−0·22
4680	−29·50	−0·58	4640	−29·50	0·05

TABLE VI—(*continued*)

λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$
4602	−30·39	−0·33	4565	−30·68	−0·02
4554	−31·39	−0·58	4463	−32·12	0·16
4390	−33·60	−0·08	4358	−34·10	0·02
4316	−35·07	−0·17	4220	−36·50	0·28
4132	−37·92	0·69	4085	−39·40	0·32
4044	−40·81	−0·11	3951	−43·13	−0·08
3845	−46·09	−0·14	3759	−48·65	−0·05
3713	−49·6	0·4	3674	−51·36	0·04
3623	−53·1	0·1	3605	−53·91	0·00
3550	−56·1	−0·1	3484	−59·1	−0·3
3410	−62·1	0·0	3290	−68·0	0·2

TABLE VII—ROTATORY DISPERSION OF BENZOYL-*l*-SECONDARY BUTYLAMINE AT 20° C.

$c = 11{\cdot}900$ g./100 c.c. alcohol. $[\alpha] = -8{\cdot}637/(\lambda^2 - 0{\cdot}04231)$.

λ	$[\alpha]_{\text{obs}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$	λ	$[\alpha]_{\text{obs.}}$	$[\alpha]_{\text{obs.}} - [\alpha]_{\text{calc.}}$
6708	−21·18	0·01	6497	−22·75	−0·01
6362	−23·78	0·02	6104	−26·05	0·07
5893	−28·33	−0·03	5782	−29·54	0·01
5780	−29·57	0·03	5700	−30·52	0·03
5536	−32·77	0·02	5461	−33·67	0·09
5219	−37·39	0·09	5153	−38·44	0·22
5106	−39·55	−0·05	4934	−43·02	0·38
4811	−45·56	0·00	4722	−47·73	0·07
4680	−49·03	−0·12	4602	−51·06	−0·08
4554	−52·25	0·00	4358	−58·35	0·15
		"Photographic"			
4532	−52·79	0·10	4427	−55·87	0·28
4355	−58·35	0·16	4307	−60·04	0·19
4273	−61·46	0·04	4184	−65·10	−0·09
4125	−67·35	0·15	4078	−69·75	−0·15
4020	−72·40	−0·09	3967	−75·06	−0·06
3911	−77·86	0·15	3857	−80·52	0·49
3852	−80·67	0·63	3821	−83·46	−0·21
3810	−83·95	0·05	3776	−85·98	0·07
3762	−87·40	−0·39	3730	−90·77	−0·57
3659	−94·13	0·12	3618	−97·50	0·00
3566	−101·7	0·1	3508	−106·8	0·0
3451	−112·1	0·3	3404	−117·2	0·0
3341	−124·2	0·2	3290	−131·3	−0·3

TABLE VIII—ROTATORY DISPERSION OF DIETHYL-*d*-SECONDARY BUTYLAMINE AT 20° C.

Density = 0·753. Length of tube = 0·4953 dm. $[\alpha] = 26\cdot44/(\lambda^2 - 0\cdot0397)$.

λ	$[\alpha]_{obs.}$	$[\alpha]_{calc} - [\alpha]_{obs.}$	λ	$[\alpha]_{obs.}$	$[\alpha]_{calc.} - [\alpha]_{obs.}$
6438	70·5	±0·0	4021	217·7	−0·7
5893	85·8	0·3	3906	234·1	0·3
5780	89·7	0·1	3756	261·4	−0·2
5461	102·1	0·1	3623	288·6	0·6
5086	120·6	0·2	3514	316·0	±0·0
4937	130·4	−0·7	3420	343·2	−0·8
4800	138·3	0·4	3333	370·4	0·2
4735	144·0	−0·7	3257	397·8	0·7
4678	147·9	−0·3	3193	425·1	±0·0
4530	160·4	−0·5	3134	452·4	−0·6
4360	176·3	−0·4	3081	479·7	−0·7
4358	176·0	±0·0	3025	507·0	4·0
4164	198·6	−0·6	2978	534·3	5·7

TABLE IX—ROTATORY DISPERSION OF DIETHYL-*d*-SECONDARY BUTYLAMINE HYDROCHLORIDE

In water at 20° C. $c = 6\cdot29$ g./100 c.c. $l = 6$ dm.
$[\alpha] = 1\cdot375/(\lambda^2 - 0\cdot0225)$.

λ	$[\alpha]_{obs.}$	$[\alpha]_{calc.} - [\alpha]_{obs.}$	λ	$[\alpha]_{obs.}$	$[\alpha]_{calc.} - [\alpha]_{obs.}$
6708	3·42	−0·20	4680	6·95	0·05
6362	3·76	−0·16	4602	7·53	−0·26
6104	3·92	0·01	4565	7·45	−0·04
5893	4·24	±0·00	4427	7·98	−0·04
5782	4·42	−0·01	4358	8·04	0·17
5780	4·48	−0·06	4348	8·51	−0·25
5700	4·51	0·04	4206	9·04	−0·13
5461	4·96	0·04	4105	9·56	−0·14
5218	5·44	0·06	4005	10·10	−0·11
5153	5·70	−0·04	3896	10·62	0·02
5086	5·83	±0·00	3821	11·15	±0·00
4936	6·39	−0·17	3741	11·69	0·01
4811	6·63	−0·05	3672	12·21	0·04
4800	6·60	0·01	3615	12·73	−0·02
4722	6·97	−0·11	3530	13·29	0·16
4707	6·92	−0·01			

SUMMARY

The rotatory dispersion of secondary butylamine is largely due to circular dichroism in the absorption bands at 2200 A associated with a lone-pair of

electrons on the amino-nitrogen atom, whereas the dissymmetry of a spirane nucleus is insufficient to produce activation of these bands.

Secondary butylamine shows anomalous rotatory dispersion for which a satisfactory explanation is put forward.

The rotatory dispersion of acyl-butylamines, butyl-urea, and of butyl-ammonium hydrochloride is simple. The activation of the lone-pair occurs only in the free base.

A new theory to account for the observed changes of rotatory power of amino bases on neutralization is put forward.

The absolute configuration of laevo-secondary butylamine is defined.

Saturation of the lone-pair of the tertiary base, diethyl-butylamine, causes the active wave-length from the Drude dispersion equations to move further into the ultra-violet, thus confirming the theory.

Lowry's results for the nicotine series are reinterpreted according to the present theory of activation of the absorption bands associated with the lone-pair.

REFERENCES

Bielecki and Henri 1913 *C.R. Acad. Sci., Paris*, **156**, 1860.
Boys 1934 *Proc. Roy. Soc.* A, **144**, 655.
Fischer and Kametaka 1909 *Liebigs Ann.* **365**, 7.
Hartley and Dobbie 1900 *J. Chem. Soc.* **77**, 318.
Leithe 1930 *Ber. dtsch. chem. Ges.* **63**, 804.
Levene and Marker 1936 *J. Biol. Chem.* **115**, 267.
Ley and Arends 1932 *Z. phys. Chem.* B, **17**, 177.
Ley and Ulrich 1909 *Ber. dtsch. chem. Ges.* **42**, 3440.
Mulliken 1935 *J. Chem. Phys.* p. 507.
Pope and Gibson 1912 *J. Chem. Soc.* **101**, 1702.
Rule 1930 *Trans. Faraday Soc.*, "Optical Rotatory Power", p. 321.
Urban 1904 *Chem. Zbl.* **1**, 997.

Rotatory Dispersion in the Amine Series
III—The Absorption Spectra of Diamines

By W. C. G. Baldwin, M.A., Ph.D.*

Laboratory of Physical Chemistry, Cambridge University

(*Communicated by R. G. W. Norrish, F.R.S.—Received* 21 *May* 1937)

In view of the striking effects on the rotatory powers of optically active amines which occur when the lone-pair of electrons on the amino-nitrogen atom is saturated, as described in the two preceding papers, it was of importance to investigate the corresponding changes which are indicated in the absorption spectra of simple amino compounds. Absorption spectra are therefore now recorded for hydrazine and for ethylene-diamine, and the effect on these spectra of varying the acid concentration of the solutions has been measured.

No previous measurements of the absorption of hydrazine in solution have been recorded, but Sunao and Imanishi (1931) observed a series of bands between 2400 and 2200 A in the vapour. Ethylene-diamine has been examined by Bielecki and Henri (1913), and Castille and Ruppol (1928) have investigated its hydrochloride. The latter authors reported a weak absorption band in aqueous solutions of the salt ($\epsilon_{max} = 0{\cdot}4$ at 2631 A), but this observation is not confirmed by the present results. It may be attributed to the presence of an impurity, since, in the first series of measurements now described, made with material of commercial purity, a similar band ($\epsilon_{max} = 2$ at 2631 A) was also seen, but after recrystallization of the salt it disappeared completely.

I—HYDRAZINE

(*a*) *Pure Base*—The pure base begins to absorb appreciably at 2500–2600 A, and, in solution in water, at 2300–2400 A. This shift is normal and may be attributed to interaction with the solvent. The presence of water, and thus the formation of a hydrate, does not completely suppress the absorption due to the lone-pairs on the nitrogen atoms, although it appears to produce a small decrease in the intensity of absorption. The present observations with amines indicate that a pure anhydrous base will always absorb more strongly than its aqueous solution, and the spectral "cut-off" usually occurs at about 200 A nearer to the visible region (see fig. 1).

---

* Christ's College Research Scholar.

(*b*) *Hydrazine Dihydrochloride*—A solution prepared by adding acid to the hydrazine, in the proportion of two molecules to one of base, showed no appreciable absorption at wave-lengths longer than 2070 A, when measured against a solution containing acid of equivalent concentration, in order to confine any measured effect entirely to the hydrazinium ion. The limiting value for the molecular extinction coefficient is $\epsilon < 0{\cdot}09$ at 2070 A. This result is to be expected when the nitrogen lone-pairs are completely saturated.

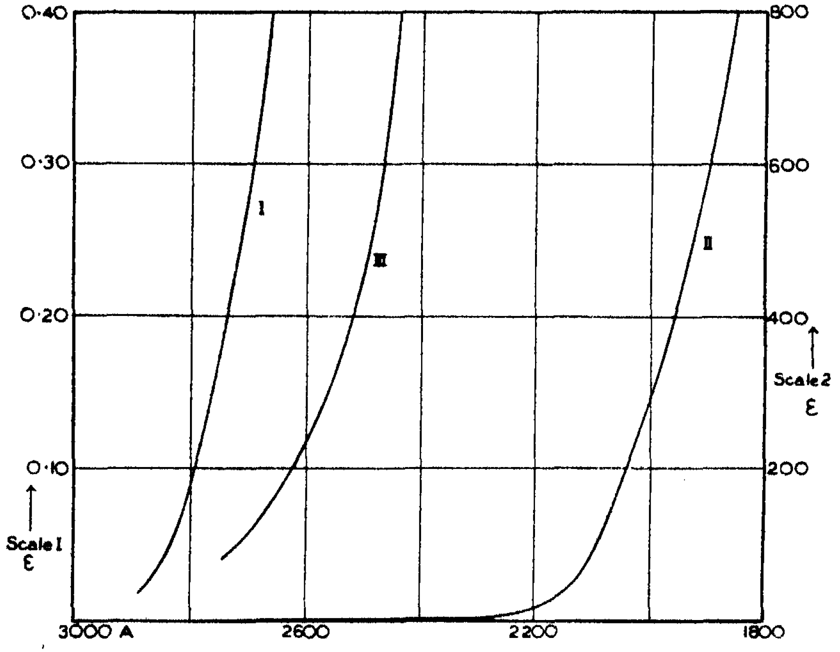


Fig. 1—Curve I. Absorption spectrum of the anhydrous base on scale 1; Curve II. Absorption spectrum of hydrazine in water, on scale 2; Curve III. Absorption spectrum of hydrazine in water on scale 1.

(*c*) *Hydrazine Monohydrochloride*—A solution of this substance, also compared with an acid "blank", showed no appreciable absorption at those wave-lengths at which the absorption of the free base is observed ($\epsilon < 0{\cdot}29$ at 2050 A, cf. $\epsilon = 186$ for the base). This remarkable result suggests that no normal lone-pair is present in the mono-salt. Thus, it may be concluded that whilst the first amino group is saturated by the proton from the acid molecule, the second must become saturated either through co-ordination with water, or, more probably, through an interaction between the residual lone-pair and the positive $\overset{+}{N}H_3$ group at the other end of the molecule, so that their characteristic absorption is no longer manifested. This explanation is in harmony with the well-known chemical fact that hydrazine, being nearly ten times weaker as a base than ammonia, is effectively only a monacid base in solution, and no great difference between the properties of the mono- and di-salts of hydrazine in water is to be expected since the latter are invariably hydrolysed. A simple displacement of the bands of the lone-pair towards shorter wave-lengths might, however, be expected, in view of the proximity of the lone-pair on one nitrogen atom to the positive charge on the other.

(*d*) *Co-ordination of Hydrazine*—Attempts were made to saturate the lone-pairs in a different way, e.g. by forming co-ordination compounds of hydrazine. Thus, the zinc chloride compound was prepared, but in common with others, it proved to be unstable in solution.

II—ETHYLENE-DIAMINE

(*a*) *Pure Base*—The purified base in aqueous solution gave an absorption curve rising steeply beyond 2350 A to $\epsilon = 81$ at 2150 A, in agreement with the limited observations of Bielecki and Henri (1913).

(*b*) *Ethylene-diamine Monochloride*—A solution containing unimolecular proportions of caustic soda and ethylene-diamine dihydrochloride gave the absorption curve II in fig. 2. There is no similarity between the absorption spectra of the mono-salts of hydrazine and of ethylene-diamine in water. In the former case, saturation seems to be complete, whereas in the latter, the curve is intermediate between those for the base and for the salt. Ethylene-diamine therefore behaves optically as a normal diacidic base, in harmony with the fact that its basic strength is considerably higher than that of hydrazine. The difference between the absorption spectra of the mono-salts of the two bases would be expected if the curve for hydrazine monohydrochloride is displaced towards shorter wave-lengths as a result of interaction between the lone-pair on one nitrogen atom and the positive charge on the other, since in the case of the ethylene-diamine ion such interaction must be considerably reduced on account of the greater separation between the two nitrogen atoms.

(*c*) *Ethylene-diamine Dihydrochloride*—The di-salt shows no appreciable absorption in aqueous solution (curve III, fig. 2).

Conclusion—These results seem to indicate conclusively that the absorption band in the region of 2200 A observed for simple amines is characteristic of the lone-pair on the nitrogen atom, and can be suppressed progressively as the basicity of the solutions is decreased. Results obtained by Castille and Ruppol (1928) for cadaverine, putresceïne and guanidine show that the same generalization is true also for more complicated amines. Very weak absorption bands at longer wave-lengths observed by them with putresceïne dihydrochloride and guanidine hydrochloride may be attributed to impurities, since corresponding bands now observed could be eliminated by rigorous purification.

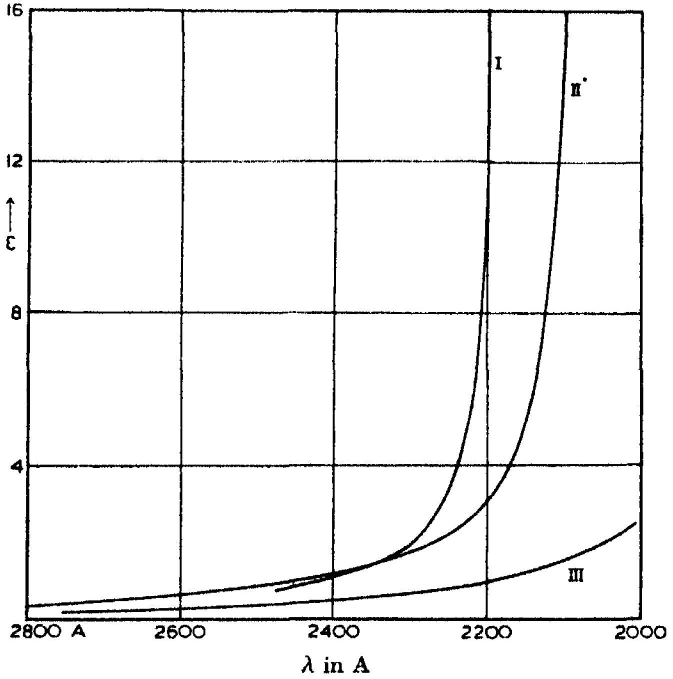


Fig. 2—Curve I. Absorption spectrum of ethylene-diamine in water; Curve II. Absorption spectrum of ethylene-diamine monohydrochloride in water; Curve III. Absorption spectrum of ethylene-diamine dihydrochloride in water.

Materials—Hydrazine hydrate was distilled twice from solid caustic soda, and once from a large quantity of barium oxide in the absence of air. B.P. 113·5° C.

Ethylene-diamine was kindly supplied by Dr. F. G. Mann, to whom the author expresses his thanks. He also wishes to thank Dr. C. B. Allsopp for help and advice.

SUMMARY

Absorption spectra are recorded for aqueous solutions of hydrazine and ethylene-diamine, and of their mono- and di-hydrochlorides. The suppression of the absorption due to the nitrogen lone-pairs of the bases, as these are progressively saturated, is discussed.

REFERENCES

Bielecki and Henri 1913 *C.R. Acad. Sci., Paris*, **156**, 1860.
Castille and Ruppol 1928 *Bull. Soc. Chim. biol., Paris*, **10**, 623.
Sunao and Imanishi 1931 *Chem. Zbl.* **1**, 3330.

Solution of Oseen's Equations for an Inclined Elliptic Cylinder in a Viscous Fluid

BY D. MEKSYN, D.SC., *Imperial College of Science and Technology*

(*Communicated by L. Bairstow, F.R.S.—Received* 21 *May* 1937)

1—INTRODUCTION

Solutions of Oseen's equations have previously been obtained for the circular cylinder (Lamb 1916; Bairstow, Cave and Lang 1923; Southwell and Squire 1933), circular and elliptic cylinders (Berry and Swain 1923; Harrison 1924; Faxén 1927) and flat plate (Berry and Swain 1923; Piercy and Winny 1933).

After eliminating the pressure from the equations of steady motion of a viscous fluid in two dimensions the equations become

$$\nu\nabla^2\zeta = u\frac{\partial\zeta}{\partial x}+v\frac{\partial\zeta}{\partial y}. \tag{1·1}$$

Oseen's modification is

$$\nu\nabla^2\zeta = V\frac{\partial\zeta}{\partial x}, \tag{1·2}$$

where V is the undisturbed velocity of the fluid and the axis of x is in the direction of the undisturbed motion.

Boussinesq, and later Southwell and Squire, showed that equation (1·1) can be converted to equation (1·2) if u and v are harmonic functions satisfying the relations

$$\left.\begin{aligned} u &= -\frac{\partial\alpha}{\partial x} = -\frac{\partial\beta}{\partial y}, \\ v &= -\frac{\partial\alpha}{\partial y} = \frac{\partial\beta}{\partial x}, \end{aligned}\right\} \tag{1·3}$$

and instead of x and y the variables are changed to α and β.

Most of the solutions have been obtained with $\alpha = -Vx$, $\beta = -Vy$ and with a combination of analytical and graphical methods. Southwell and Squire, however, solved by graphics the case of flow round a circular cylinder when

$$\alpha + i\beta = -V\left(z + \frac{a^2}{z}\right). \tag{1·4}$$

The method adopted in several of these solutions was that of adding a number of separate solutions of equation (1·2) and combining them so as to satisfy the boundary conditions at a finite number of points.

The present paper shows that for $\alpha + i\beta = -Vz$ all previous solutions can be obtained completely in terms of Mathieu functions; at the same time an extension is obtained which covers the general case of the inclined elliptic cylinder.

The forces on the cylinder have been evaluated, and for very small values of the Reynolds number ($\not> 4$) the lift and drag curves as functions of angle of incidence are found to be similar to those of a normal aerofoil, though no experiments are available at such small Reynolds numbers. For a Reynolds number of 3·3 the curves for lift, drag, etc., are shown in § 7.

The drag is much greater, whereas the lift is little different from that of an ordinary aerofoil. The curves for drag, lift and ratio lift/drag as functions of the angle of incidence are similar to those of an ordinary aerofoil, although the corresponding positions of maxima and minima are different.

Lift rapidly increases with increase of angle of incidence, reaches its maximum value at 45°, and then decreases to zero.

Drag increases more slowly, so that the curve lift/drag passes through a maximum.

Method of Solution—It is assumed initially that the motion is irrotational and then use is made of the corresponding velocity potential in the equation of motion of vorticity in a viscous fluid.

The axes of reference have been taken as the principal axes of the ellipse and not along and normal to the undisturbed stream.

2—IRROTATIONAL MOTION OF FLUID

Consider an elliptic cylinder at rest in an infinite non-viscous fluid.

The x and y axes (fig. 1) are along the major and minor axes of the cylinder. The undisturbed fluid moves with velocity V at an angle θ with the positive direction of x.

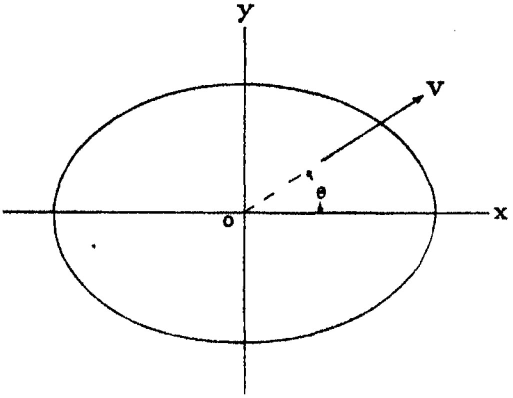


FIG. 1

The velocity potential and the stream functions of the undisturbed motion are

$$\left.\begin{aligned}\alpha &= -V\cos\theta\,.\,x - V\sin\theta\,.\,y,\\ \beta &= -V\cos\theta\,.\,y + V\sin\theta\,.\,x.\end{aligned}\right\} \tag{2·1}$$

Introducing elliptic co-ordinates

$$x = c\cosh\xi\,\cos\eta,\quad y = c\sinh\xi\,\sin\eta, \tag{2·2}$$

we obtain for α and β the expressions

$$\left.\begin{aligned}\alpha &= -Vc[\cosh\xi\,\cos\eta\,\cos\theta + \sinh\xi\,\sin\eta\,\sin\theta],\\ \beta &= -Vc[\sinh\xi\,\sin\eta\,\cos\theta - \cosh\xi\,\cos\eta\,\sin\theta].\end{aligned}\right\} \tag{2·3}$$

The equation to the ellipse is

$$\left.\begin{aligned}\frac{x^2}{c^2\cosh^2\xi} + \frac{y^2}{c^2\sinh^2\xi} = 1,\\ \xi = \xi_0.\end{aligned}\right\} \tag{2·4}$$

The irrotational motion of the fluid round the elliptic cylinder is defined by the potential and stream functions

$$\left.\begin{aligned}\phi_0 &= -Vce^{\xi_0}\cosh(\xi-\xi_0)\cos(\eta-\theta),\\ \psi_0 &= -Vce^{\xi_0}\sinh(\xi-\xi_0)\sin(\eta-\theta).\end{aligned}\right\} \tag{2·5}$$

It is easy to verify that (2·5) is the required solution. The condition $\xi = \xi_0$ makes ψ_0, the stream function, vanish; also the component velocities at infinity have the required values (Ramsay 1920, p. 111) $V\cos\theta$ and $V\sin\theta$.

3—Rotational Motion

(*a*)—*Equation of Vorticity*

The motion of a fluid round a cylinder is considered as though produced in two stages:

(i) At the beginning there is an irrotational motion defined by the expressions (2·5) for the velocity potential and stream function.

(ii) This motion is disturbed by friction between the boundary of the solid body and the fluid which gives rise to a layer of vortices; and our task is to find the motion of these vortices in a viscous fluid.

We have
$$u = -\frac{\partial\psi}{\partial y}, \quad v = \frac{\partial\psi}{\partial x}. \tag{3·1}$$

The vorticity is
$$\zeta = \frac{\partial v}{\partial x}-\frac{\partial u}{\partial y} = \nabla^2\psi. \tag{3·2}$$

The equation of motion in terms of vorticity is given by

$$\frac{\partial\zeta}{\partial t}+u\frac{\partial\zeta}{\partial x}+v\frac{\partial\zeta}{\partial y} = \nu\nabla^2\zeta, \tag{3·3}$$

where
$$\nu = \frac{\mu}{\rho},$$

μ being the coefficient of viscosity, and ρ the density of the fluid.

This equation is solved by successive approximations, inserting in the first approximation the values for u and v as obtained from the undisturbed motion; or

$$\frac{\partial\zeta}{\partial t}-\frac{\partial\alpha}{\partial x}\frac{\partial\zeta}{\partial x}-\frac{\partial\alpha}{\partial y}\frac{\partial\zeta}{\partial y} = \nu\nabla^2\zeta. \tag{3·4}$$

The equation (3·4) expressed in elliptic co-ordinates ξ, η becomes

$$\left.\begin{aligned}\frac{1}{h^2}\frac{\partial\zeta}{\partial t}-\left(\frac{\partial\alpha}{\partial\xi}\frac{\partial\zeta}{\partial\xi}+\frac{\partial\alpha}{\partial\eta}\frac{\partial\zeta}{\partial\eta}\right)&=\nu\left(\frac{\partial^2\zeta}{\partial\xi^2}+\frac{\partial^2\zeta}{\partial\eta^2}\right),\\ h^2&=\frac{1}{c^2(\cosh^2\xi-\cos^2\eta)}.\end{aligned}\right\}\tag{3·5}$$

Assume a solution
$$\zeta = e^{-\frac{\alpha}{2\nu}}f'(\xi,\eta,t).\tag{3·6}$$

Equation (3·5) is thus transformed, after dropping the factor $\exp\left(-\frac{\alpha}{2\nu}\right)$, and making use of (2·3), into

$$\frac{1}{h^2}\frac{\partial f'}{\partial t}+\frac{1}{4\nu}\left[\left(\frac{\partial\alpha}{\partial\xi}\right)^2+\left(\frac{\partial\alpha}{\partial\eta}\right)^2\right]f'=\nu\nabla^2f'.\tag{3·7}$$

If
$$f'(\xi,\eta,t)=e^{-\lambda t}f_1(\xi)f_2(\eta),\tag{3·8}$$

we obtain the following equations for f_1 and f_2:

$$\left.\begin{aligned}\frac{\partial^2f_1}{\partial\xi^2}+\left[\frac{\lambda c^2}{\nu}\cosh^2\xi-\frac{V^2c^2}{4\nu^2}\cosh^2\xi-E\right]f_1&=0,\\ \frac{\partial^2f_2}{\partial\eta^2}+\left[-\frac{\lambda c^2}{\nu}\cos^2\eta+\frac{V^2c^2}{4\nu^2}\cos^2\eta+E\right]f_2&=0,\end{aligned}\right\}\tag{3·9}$$

where E is a constant which is found from the condition that $f_2(\eta)$ is a periodic function in η.

For steady motion $\lambda=0$, and the solutions for f_1 and f_2 are Mathieu functions.

Let
$$\frac{Vc}{2\nu}=k,\quad E+\frac{k^2}{2}=4a,\quad k^2=32q.\tag{3·10}$$

Then, after an easy transformation, we obtain

$$\frac{\partial^2f_2}{\partial\eta^2}+[4a+16q\cos2\eta]f_2=0,\tag{3·11a}$$

$$\frac{\partial^2f_1}{\partial\xi^2}-[4a+16q\cosh2\xi]f_1=0.\tag{3·11b}$$

(*b*)—*Solution of the equations*

The solutions of (3·11) are Mathieu ordinary and associated functions. We denote (Whittaker and Watson 1927, p. 409) by se_{2n}, se_{2n+1}, ce_0, ce_{2n},

ce_{2n+1} Mathieu functions which reduce to $\sin 2n\eta$, $\sin(2n+1)\eta$, 1, $\cos 2n\eta$ and $\cos(2n+1)\eta$, when $q = 0$.

We have now to construct the solution of equation (3·11*b*), the associated Mathieu functions.

Since the expression for ζ includes an exponential factor

$$\exp\left\{\frac{Vc}{2\nu}(\cosh\xi\cos\eta\cos\theta+\sinh\xi\sin\eta\sin\theta)\right\},$$

which tends to infinity when $\xi\to\infty$, the associated Mathieu functions must vanish exponentially at infinity, in order that ζ shall vanish at infinity.

Goldstein (1927) has shown that Mathieu functions can be expanded in series of Bessel functions, the coefficients being proportional to the corresponding terms in Fourier's expansion.

We try, accordingly, to expand the associated Mathieu functions in series of Bessel functions K_n which vanish at infinity (Gray and Mathews 1922, Chap. III).

An example of such an expansion is given, and then only the final results are quoted.

Write Mathieu's equation in the form

$$\frac{d^2y_1}{dx^2}+[4a+16q\cos 2x]\,y_1 = 0, \tag{3·12}$$

and the associated equation as

$$\frac{d^2y_2}{dx^2}-[4a+16q\cosh 2x]\,y_2 = 0, \tag{3·13}$$

where

$$k^2 = 32q. \tag{3·14}$$

Let the solution of (3·12) be

$$\left.\begin{aligned} ce_{2n} &= \sum_{r=0}^{\infty} A_{2r}\cos 2rx, & ce_{2n+1} &= \sum_{r=0}^{\infty} A_{2r+1}\cos(2r+1)x,\\ se_{2n} &= \sum_{r=1}^{\infty} B_{2r}\sin 2rx, & se_{2n+1} &= \sum_{r=0}^{\infty} B_{2r+1}\sin(2r+1)x.\end{aligned}\right\} \tag{3·15}$$

Inserting the expression for ce_{2n} into (3·12) we find that the coefficients A_{2r} satisfy the following recurrence relations:

$$\left.\begin{aligned} aA_0+2qA_2 &= 0,\\ 4qA_0+(a-1)A_2+2qA_4 &= 0,\\ 2qA_{2r-2}+(a-r^2)A_{2r}+2qA_{2r+2} &= 0, \quad r\geqslant 2.\end{aligned}\right\} \tag{3·16}$$

The corresponding expression for the associated Mathieu function Cek_{2n} is found by introducing a new variable

$$k\cosh x = z, \tag{3·17}$$

when equation (3·13) becomes

$$(z^2-k^2)\frac{d^2y_2}{dz^2}+z\frac{dy_2}{dz}-[z^2+4a-16q]\,y_2 = 0. \tag{3·18}$$

Assuming an expansion

$$y_2 = Cek_{2n} = \sum_{r=0}^{\infty} \bar{A}_{2r}\, K_{2r}(z), \tag{3·19}$$

we substitute for Cek_{2n} in (3·18) and make use of the equations

$$\left.\begin{aligned} z^2K_r''+zK_r'-(r^2+z^2)\,K_r &= 0,\\ K_r'' &= \tfrac{1}{4}(K_{r-2}+2K_r+K_{r+2}),\\ K_0'' &= \tfrac{1}{2}(K_2+K_0), \end{aligned}\right\} \tag{3·20}$$

where dashes to K denote differentiation with respect to z (Gray and Mathews 1922, pp. 20–2).

Equating to zero the coefficients of the K_r's we find that the $\bar{A}$'s satisfy the same relations (3·16) as the A's. Hence

$$\bar{A}_{2r} = A_{2r}. \tag{3·21}$$

The expansions for Sek_{2n} and Sek_{2n+1} are more complicated.

Assume a solution (Goldstein 1927, p. 311)

$$y_2 = u\tanh x, \tag{3·22}$$

where

$$u = \sum_{r=0}^{\infty}(2r+1)\,\bar{B}_{2r+1}K_{2r+1}(z). \tag{3·23}$$

The corresponding equation for u is then

$$(z^2-k^2)\,u''+zu'-(z^2+4a-16q)\,u+2k^2\frac{d}{dz}\left(\frac{u}{z}\right) = 0, \tag{3·24}$$

and making use, in addition to (3·20), of the identity

$$4r\frac{d}{dz}\left(\frac{K_r}{z}\right) = K_{r-2}-K_{r+2}, \tag{3·25}$$

we find that the $\bar{B}_{2r+1}$'s satisfy the same relations as the B_{2r+1}'s in the corresponding Mathieu functions.

Finally, the following expressions for the associated Mathieu functions are obtained:

$$\left.\begin{aligned} Cek_{2n}(x) &= \sum_{r=0}^{\infty} A_{2r}^{(2n)} K_{2r}(k\cosh x), \\ Cek_{2n+1}(x) &= \sum_{r=0}^{\infty} A_{2r+1}^{(2n+1)} K_{2r+1}(k\cosh x), \\ Sek_{2n}(x) &= \tanh x \sum_{r=1}^{\infty} 2r B_{2r}^{(2n)} K_{2r}(k\cosh x), \\ Sek_{2n+1}(x) &= \tanh x \sum_{r=0}^{\infty} (2r+1)\, B_{2r+1}^{(2n+1)} K_{2r+1}(k\cosh x). \end{aligned}\right\} \qquad (3{\cdot}26)$$

In our case $x = \xi.$

Using the asymptotic expression for the K_n functions, i.e.

$$K_n(x) = \sqrt{\frac{\pi}{2x}}\, e^{-x}\left\{1 + \frac{4n^2-1^2}{1!\,8x} + \ldots\right\},$$

where $x = k\cosh\xi$

or

$$\left.\begin{aligned} K_n &= \sqrt{\frac{\pi}{2k\cosh\xi}}\, e^{-\kappa\cosh\xi}\{1+\ldots\}, \\ k &= \frac{Vc}{2\nu}, \end{aligned}\right\} \qquad (3{\cdot}27)$$

which will reduce the exponential factor

$$\exp\left\{\frac{Vc}{2\nu}(\cosh\xi\,\cos\eta\,\cos\theta + \sinh\xi\,\sin\eta\,\sin\theta)\right\}$$

to zero at infinity in ζ.

Finally

$$\zeta = e^{\frac{Vc}{2\nu}(\cosh\xi\,\cos\eta\,\cos\theta + \sinh\xi\,\sin\eta\,\sin\theta)} \times \left\{\sum_{n=0}^{\infty} C_n\, Cek_n(\xi)\, ce_n(\eta) + \sum_{n=1}^{\infty} S_n\, Sek_n(\xi)\, se_n(\eta)\right\}, \qquad (3{\cdot}28)$$

where C_n's and S_n's are coefficients to be found from the boundary conditions.

4—STREAM FUNCTION OF ROTATIONAL MOTION

(a) *Green's Function*

It is necessary now to find the corresponding stream function, i.e. to solve the equation

$$\frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}=h^2\left(\frac{\partial^2\psi}{\partial \xi^2}+\frac{\partial^2\psi}{\partial \eta^2}\right)=\zeta. \tag{4·1}$$

Instead of solving (4·1) in the x, y plane consider the plane ξ, η with the corresponding vorticity as ζ/h_2; the region of integration is now not the plane x, y external to the ellipse, but a strip of plane ξ, η bounded by lines

$$\eta=0, \quad \eta=2\pi, \quad \xi=\xi_0.$$

In order to solve equation (4·1) we have to find Green's function $G(\xi',\eta';\xi,\eta)$, where ξ',η' and ξ,η are corresponding source and field points (fig. 2).

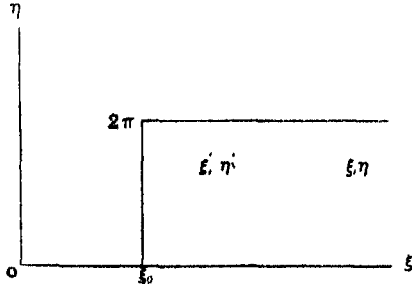


FIG. 2

As is known the solution is based on Green's formula

$$\iint(u\nabla^2v-v\nabla^2u)\,dx\,dy=\int\left(u\frac{\partial v}{\partial n}-v\frac{\partial u}{\partial n}\right)ds, \tag{4·2}$$

where u and v are functions of x and y, and the differentiation $\dfrac{\partial}{\partial n}$ is taken along an external normal.

If v satisfies an equation similar to ψ (4·1) we insert in (4·2) instead of ∇^2v its given value, and instead of u the corresponding Green's function; and, thus, obtain in the well-known way the expression for v.

One of the conditions that Green's function is required to satisfy is that the integral

$$\int u\frac{\partial v}{\partial n}\,ds \tag{4·2a}$$

taken round the boundary vanishes; this is usually achieved by making u vanish on the boundary.

Such a procedure, however, cannot be applied in our case, since it would make $\eta = 0$ and $\eta = 2\pi$ or $y = 0$ a stream line, which would correspond to a case when there is an infinitely thin, infinitely long screen along the axis of x.

We can, however, make the integral (4·2a) vanish if u vanishes along $\xi = \xi_0$ and if u and v are periodic functions in η with a period 2π. The integrand (4·2a) along the line $\xi = \xi_0$ from $\eta = 0$ to $\eta = 2\pi$ will vanish, and the integrands along the lines $\eta = 0$ and $\eta = 2\pi$ will mutually cancel out; for, when the integrand is periodic in η with period 2π, $\frac{\partial v}{\partial n}$ has opposite signs at the corresponding points, since the normal is always an external one.

Now, Green's function for an infinite half-plane, bounded at one side by the line $\xi = \xi_0$, is

$$\frac{1}{4\pi}\log\frac{(\xi-\xi')^2+(\eta-\eta')^2}{(\xi+\xi'-2\xi_0)^2+(\eta-\eta')^2}.$$

It vanishes along the line $\xi = \xi_0$.

The function sought, however, must be periodic in η with period 2π; so that the expression for Green's function will be

$$G(\xi',\eta';\xi,\eta) = \frac{1}{4\pi}\sum_{n=-\infty}^{n=+\infty}\log\frac{(\xi-\xi')^2+(\eta-\eta'-2\pi n)^2}{(\xi+\xi'-2\xi_0)^2+(\eta-\eta'-2\pi n)^2}. \tag{4·3}$$

This sum, as in the case of a rectangular parallelepiped, can be expressed as a definite integral (Courant and Hilbert 1931).

Consider the integral

$$\int_0^\infty \frac{e^{-at}-e^{-bt}}{t}\,dt = \log\frac{b}{a}, \tag{4·4}$$

where $a > 0$ and $b > 0$.

Combining (4·3) and (4·4) we easily find

$$G = \frac{1}{4\pi}\int_0^\infty \frac{e^{-(\xi+\xi'-2\xi_0)^2 t}-e^{-(\xi-\xi')^2 t}}{t}\sum_{n=-\infty}^{\infty} e^{-(\eta-\eta'-2\pi n)^2 t}\,dt, \tag{4·5}$$

where it has been assumed that the integration and the summation can be interchanged.

The infinite sum can be expressed in θ functions (Whittaker and Watson 1927, p. 464)

$$\theta_3(z,q) = \sum_{n=-\infty}^{\infty} q^{n^2}e^{2izn}, \quad q = e^{i\pi\tau}. \tag{4·6}$$

It is easily seen that in our case

$$q = e^{-4\pi^2 t}, \quad z = \frac{2\pi(\eta-\eta')\,t}{i}, \tag{4·7}$$

and accordingly

$$\mathrm{G} = \frac{1}{4\pi}\int_0^\infty \frac{e^{-[(\xi+\xi'-2\xi_0)^2+(\eta-\eta')^2]\,t} - e^{-[(\xi-\xi')^2+(\eta-\eta')^2]\,t}}{t}\,\theta_3(-2\pi i(\eta-\eta')\,t,\, e^{-4\pi^2 t})\,dt. \tag{4·8}$$

Next, apply Jacobi's transformation to the θ function, namely (Whittaker and Watson 1927, p. 475)

$$\theta_3(z/\tau) = (-i\tau)^{-\frac{1}{2}}\exp\left(\frac{z^2}{i\pi\tau}\right)\theta_3\left(\frac{z}{\tau}\Big/-\frac{1}{\tau}\right). \tag{4·9}$$

In our case $\tau = 4\pi i t,$ (4·10)

whence

$$\theta_3(-2\pi i(\eta-\eta')\,t/4\pi i t) = (4\pi t)^{-\frac{1}{2}}\exp\left(\frac{4\pi^2(\eta-\eta')^2\,t^2}{4\pi^2 t}\right)\theta_3\left(-\frac{(\eta-\eta')}{2}\Big/-\frac{1}{4\pi i t}\right)$$

$$= \frac{1}{\sqrt{(4\pi t)}}\,e^{(\eta-\eta')^2 t}\sum_{n=-\infty}^{\infty} e^{-in(\eta-\eta')-\frac{n^2}{4t}}, \tag{4·11}$$

and

$$\mathrm{G} = \frac{1}{8\pi^{\frac{1}{2}}}\int_0^\infty \frac{e^{-(\xi+\xi'-2\xi_0)^2 t} - e^{-(\xi-\xi')^2 t}}{t^{\frac{1}{2}}}\left[1+2\sum_{n=1}^{\infty} e^{-\frac{n^2}{4t}}\cos n(\eta-\eta')\right]dt. \tag{4·12}$$

Put $t = \frac{1}{t'^2}$

to find that

$$\mathrm{G}(\xi',\eta';\xi,\eta) = \frac{1}{4\pi^{\frac{1}{2}}}\int_0^\infty\left[e^{-\frac{(\xi+\xi'-2\xi_0)^2}{t^2}} - e^{-\frac{(\xi-\xi')^2}{t^2}}\right]\left[1+2\sum_{n=1}^{\infty} e^{-\frac{n^2 t^2}{4}}\cos n(\eta-\eta')\right]dt, \tag{4·13}$$

where the dashes to the t's have been dropped.

The stream function is

$$\psi(\xi,\eta) = \int_{\xi'=\xi_0}^{\infty}\int_{\eta'=0}^{2\pi}\zeta(\xi',\eta')\,\mathrm{G}(\xi',\eta';\xi,\eta)\frac{d\xi'\,d\eta'}{h^2}, \tag{4·14}$$

where $\frac{1}{h^2} = c^2(\cosh^2\xi' - \cos^2\eta').$

The ξ and η components of the velocity are

$$\left.\begin{gathered} -h\frac{\partial\psi}{\partial\eta}, \quad h\frac{\partial\psi}{\partial\xi}, \\ \xi = \xi_0. \end{gathered}\right\} \tag{4·15}$$

It can be verified that the normal flow vanishes at the boundary; i.e.

$$\left(\frac{\partial\psi}{\partial\eta}\right)_{\xi=\xi_0} = 0.$$

(*b*) *Determination of Coefficients in the Expansion of* ζ

We have now to find the values of the coefficients C_n and S_n in ζ.

Our boundary condition is that the tangential motion at the boundary shall be a prescribed function of the variable η; it is, of course, a periodic function, and can be expanded in a Fourier's series.

The tangential velocity is

$$h\frac{\partial\psi}{\partial\xi}, \quad \xi = \xi_0.$$

The problem is simplified if it is assumed that $\frac{\partial\psi}{\partial\xi}$ is the prescribed function at the boundary, i.e. at $\xi = \xi_0$.

Assume, therefore, that

$$\left(\frac{\partial\psi}{\partial\xi}\right)_{\xi=\xi_0} = E_0 + \sum_1^\infty E_n \cos n\eta + \sum_1^\infty F_n \sin n\eta, \tag{4·16}$$

$\left(\frac{\partial\psi}{\partial\xi}\right)_{\xi=\xi_0}$ is evaluated by (4·14).

From (4·13)

$$\left(\frac{\partial G}{\partial\xi}\right)_{\xi=\xi_0} = -\frac{1}{\sqrt{2}\,\pi^{\frac{1}{2}}}\int_0^\infty \sqrt{(\xi'-\xi_0)}\, e^{-\frac{1}{2}(\xi'-\xi_0)t^2}\left[1 + 2\sum_1^\infty e^{-\frac{1}{2}(\xi'-\xi_0)\frac{n^2}{t^2}} \cos n(\eta-\eta')\right]dt,$$

where the variable t has been changed into

$$t' = \frac{\sqrt{2(\xi'-\xi_0)}}{t}$$

and the dashes to the t's have been dropped.

This expression can be integrated with respect to t, since

$$\int_0^\infty e^{-\left(a^2t^2+\frac{b^2}{t^2}\right)}dt = \frac{\sqrt{\pi}}{2a}e^{-2ab}.$$

Accordingly $$\left(\frac{\partial G}{\partial \xi}\right)_{\xi=\xi_0} = -\frac{1}{2\pi}\left[1+2\sum_1^\infty e^{-n(\xi'-\xi_0)}\cos n(\eta-\eta')\right], \tag{4·17}$$

whence we find from (4·17) and (4·14)

$$\left(\frac{\partial \psi}{\partial \xi}\right)_{\xi=\xi_0} = -\frac{c^2}{4\pi}\int_{\xi'=\xi_0}^{\infty}\int_{\eta'=0}^{\infty}\zeta(\xi',\eta')\left[1+2\sum_1^\infty e^{-n(\xi'-\xi_0)}\cos n(\eta-\eta')\right] \times(\cosh 2\xi'-\cos 2\eta')\,d\xi'\,d\eta'. \tag{4·18}$$

Expressing $\cos n(\eta-\eta')$ in terms of the angles η and η', namely

$$\cos n(\eta-\eta') = \cos n\eta \cos n\eta' + \sin n\eta \sin n\eta',$$

and comparing the coefficients of the terms in $\cos n\eta$ and $\sin n\eta$ in equations (4·18) and (4·16) we obtain an infinite number of linear equations from which the coefficients C_n and S_n are to be found.

It is assumed that there is no slipping at the boundary, or

$$\frac{\partial(\psi_0+\psi)}{\partial\xi} = 0, \quad \xi = \xi_0,$$

whence $$\left(\frac{\partial \psi}{\partial \xi}\right)_{\xi=\xi_0} = -\left(\frac{\partial \psi_0}{\partial \xi}\right)_{\xi=\xi_0} = Vce^{\alpha}\sin(\eta-\theta), \tag{4·19}$$

where ψ_0 denotes the irrotational stream function (2·5).

5—Forces Acting on the Cylinder

The equations of motion are

$$\left.\begin{aligned}\frac{\partial u}{\partial t} - v\zeta &= -\frac{\partial \chi}{\partial x} + \nu\nabla^2 u,\\ \frac{\partial v}{\partial t} + u\zeta &= -\frac{\partial \chi}{\partial y} + \nu\nabla^2 v,\\ \chi &= \frac{p}{\rho} + \tfrac{1}{2}q^2 + \Omega,\end{aligned}\right\} \tag{5·1}$$

where p is the pressure, q the velocity and Ω the potential of external forces. It is assumed that $\Omega = 0$.

Now eliminating χ from (5·1) we derive the equation (3·3) for ζ in which we have substituted for u and v in the terms $u\zeta$ and $-v\zeta$, their irrotational values

$$u = -\frac{\partial\beta}{\partial y}, \quad v = \frac{\partial\beta}{\partial x}. \tag{5·2}$$

Accordingly our expression for ζ gives the condition of integration for χ.

Considering the steady motion, and substituting instead of u, v their values from (5·2) we easily transform (5·1) to our new co-ordinates ξ, η; and thus obtain

$$\left.\begin{aligned} -\frac{\partial\beta}{\partial\xi}\zeta+\nu\frac{\partial\zeta}{\partial\eta} &= -\frac{\partial\chi}{\partial\xi}, \\ -\frac{\partial\beta}{\partial\eta}\zeta-\nu\frac{\partial\zeta}{\partial\xi} &= -\frac{\partial\chi}{\partial\eta}. \end{aligned}\right\} \tag{5·3}$$

Only the expression for χ at the boundary of the ellipse, i.e. for $\xi=\xi_0$ is needed; accordingly we consider only the second equation, and find

$$\chi=\int\left(\frac{\partial\beta}{\partial\eta}\zeta+\nu\frac{\partial\zeta}{\partial\xi}\right)_{\xi=\xi_0}d\eta+\text{const.} \tag{5·4}$$

The stresses are

$$\left.\begin{aligned} p_{xx} &= -p+2\mu\frac{\partial u}{\partial x}, \\ p_{yy} &= -p+2\mu\frac{\partial v}{\partial y}, \\ p_{xy} &= \mu\left(\frac{\partial v}{\partial x}+\frac{\partial u}{\partial y}\right). \end{aligned}\right\} \tag{5·5}$$

The X component of the forces is

$$\begin{aligned} X=\int(lp_{xx}+mp_{xy})\,ds &= -\int p\frac{dy}{ds}ds+\int\left(\frac{\partial p_{xx}}{\partial x}+\frac{\partial p_{xy}}{\partial y}\right)dx\,dy \\ &= -\int p\,dy+\mu\int\nabla^2u\,dx\,dy, \end{aligned} \tag{5·6}$$

where

$$l=\frac{dy}{ds},\quad m=-\frac{dx}{ds}$$

are the direction cosines of the external normal to the cylinder.

Now

$$\nabla^2u=-\frac{\partial\zeta}{\partial y},\quad \nabla^2v=-\frac{\partial\zeta}{\partial x}. \tag{5·7}$$

Accordingly

$$X=-\int p\,dy-\mu\int\frac{\partial\zeta}{\partial y}dx\,dy=\int y\,dp+\mu\int\zeta\,dx=\int_0^{2\pi}y\frac{\partial p}{\partial\eta}d\eta+\mu\int_0^{2\pi}\zeta\frac{\partial x}{\partial\eta}d\eta, \tag{5·8}$$

where the integration is taken for $\xi=\xi_0$.

Now $$x = c\cosh\xi_0\cos\eta, \quad y = c\sinh\xi_0\sin\eta \tag{5·9}$$

and $$\frac{\partial p}{\partial\eta} = \rho\frac{\partial\chi}{\partial\eta} = \rho\left(\zeta\frac{\partial\beta}{\partial\eta} + \nu\frac{\partial\zeta}{\partial\xi}\right)_{\xi=\xi_0},$$

since we assume that $q = 0$ and $\Omega = 0$; whence

$$X = \int_0^{2\pi} c\sinh\xi_0\sin\eta\left(\rho\zeta\frac{\partial\beta}{\partial\eta} + \mu\frac{\partial\zeta}{\partial\xi}\right)d\eta - \int_0^{2\pi}\mu c\cosh\xi_0\sin\zeta\,d\eta. \tag{5·10}$$

Similarly

$$\begin{aligned} Y &= \int(lp_{xy} + mp_{yy})\,ds = \int p\,dx + \mu\int\nabla^2 v\,dx\,dy \\ &= \int p\,dx + \mu\int\frac{\partial\zeta}{\partial x}dx\,dy = -\int x\,dp + \mu\int\zeta\,dy = -\int_0^{2\pi} x\frac{\partial p}{\partial\eta}d\eta + \mu\int_0^{2\pi}\zeta\frac{\partial y}{\partial\eta}d\eta \\ &= -\int_0^{2\pi} c\cosh\xi_0\cos\eta\left[\rho\zeta\frac{\partial\beta}{\partial\eta} + \mu\frac{\partial\zeta}{\partial\xi}\right]d\eta + \mu\int_0^{2\pi} c\sinh\xi_0\sin\eta\,\zeta\,d\eta, \end{aligned}$$

where in (5·10) and (5·11) we have to put $\xi = \xi_0$. (5·11)

Consider the case $\frac{Vc}{\nu} = 1$, which corresponds to small Reynolds numbers $R < 4$ and to a very small value of q (3·10). Dropping the exponential in ζ in (3·28) and retaining only the first two terms we get

$$\zeta = CCek_1(\xi)\cos\eta + SSek_1(\xi)\sin\eta, \tag{5·12}$$

where, further, only the first terms in ce_1 and se_1 have been retained.

It is easily seen that to our degree of approximation

$$\left.\begin{aligned} \int_0^{2\pi} c\sinh\xi_0\sin\eta\,\zeta\frac{\partial\beta}{\partial\eta}d\eta &= 0, \\ \int_0^{2\pi} c\cosh\xi_0\cos\eta\,\zeta\frac{\partial\beta}{\partial\eta}d\eta &= 0. \end{aligned}\right\} \tag{5·13}$$

Finally from (5·10), (5·11), (5·12) and (5·13)

$$\left.\begin{aligned} X &= \mu c\pi S\left(\sinh\xi_0\frac{dSek_1(\xi)}{d\xi} - \cosh\xi_0 Sek_1(\xi)\right), \\ Y &= \mu c\pi C\left(-\cosh\xi_0\frac{dCek_1(\xi)}{d\xi} + \sinh\xi_0 Cek_1(\xi)\right), \\ &\xi = \xi_0. \end{aligned}\right\} \tag{5·14}$$

6—Evaluation of Vorticity

The coefficients in ζ for small values of k only are evaluated. In the example chosen

$$k = \frac{Vc}{2\nu} = \tfrac{1}{2}. \qquad (6\cdot 1)$$

It has been assumed in (5·12) that

$$\zeta = CCek_1(\xi)\cos\eta + SSek_1(\xi)\sin\eta. \qquad (6\cdot 2)$$

From (4·18) and (4·19) we obtain

$$\begin{aligned} Vce^{\xi_0}\sin(\eta-\theta) = -\frac{c^2}{4\pi}\int_0^{2\pi}\int_{\xi_0}^{\infty} & [CCek_1(\xi')\cos\eta' + SSek_1(\xi')\sin\eta'] \\ & \times\ [1+2e^{-(\xi'-\xi_0)}\cos(\eta-\eta')]\,(\cosh 2\xi' - \cos 2\eta')\,d\xi'\,d\eta'. \qquad (6\cdot 3) \end{aligned}$$

Integrating with respect to η' and comparing the terms in $\sin\eta$ and $\cos\eta$, we finally obtain

$$\left.\begin{aligned} C &= \frac{4V}{c}\,\frac{\sin\theta}{2M_2 - M_1}, & S &= -\frac{4V}{c}\,\frac{\cos\theta}{2N_2+N_1}, \\ M_1 &= \int_{\xi_0}^{\infty} Cek_1(\xi')\,e^{-\xi'}\,d\xi', & M_2 &= \int_{\xi_0}^{\infty} Cek_1(\xi')\,e^{-\xi'}\cosh 2\xi'\,d\xi', \\ N_1 &= \int_{\xi_0}^{\infty} Sek_1(\xi')\,e^{-\xi'}\,d\xi', & N_2 &= \int_{\xi_0}^{\infty} Sek_1(\xi')\,e^{-\xi'}\cosh 2\xi'\,d\xi'. \end{aligned}\right\} \qquad (6\cdot 4)$$

For our value of k (6·1), the Reynolds number is equal to

$$R = \frac{V.2a}{\nu} = \frac{2a}{c} = \frac{2a}{\sqrt{(a^2-b^2)}}. \qquad (6\cdot 5)$$

Combining (5·13), (6·1) and (6·4) we obtain finally

$$\left.\begin{aligned} X &= \frac{4\pi\rho(a+b)V^2}{(2N_2+N_1)e^{\xi_0}}\left\{-\sinh\xi_0\,\frac{dSek_1(\xi)}{d\xi} + \cosh\xi_0\,Sek_1(\xi)\right\}\cos\theta, \\ Y &= \frac{4\pi\rho(a+b)V^2}{(2M_2-M_1)e^{\xi_0}}\left\{-\cosh\xi_0\,\frac{dCek_1(\xi)}{d\xi_0} + \sinh\xi_0\,Cek_1(\xi)\right\}\sin\theta, \\ & \qquad\qquad \xi = \xi_0. \end{aligned}\right\} \qquad (6\cdot 6)$$

7—Drag and Lift

The coefficients of drag and lift are (fig. 3)

$$\left.\begin{aligned}
C_D &= \frac{X\cos\theta + Y\sin\theta}{\tfrac{1}{2}.2a\rho V^2} \\
&= \frac{4\pi(a+b)}{a(2N_2+N_1)\,e^{\xi_0}}\left[-\sinh\xi_0\frac{dSek_1(\xi)}{d\xi} + \cosh\xi_0\,Sek_1(\xi)\right]\cos^2\theta \\
&\quad + \frac{4\pi(a+b)}{a(2M_2-M_1)\,e^{\xi_0}}\left[-\cosh\xi_0\frac{dCek_1(\xi)}{d\xi} + \sinh\xi_0\,Cek_1(\xi)\right]\sin^2\theta, \\
C_L &= \frac{-X\sin\theta + Y\cos\theta}{\tfrac{1}{2}.2a\rho V^2} \\
&= \frac{a+b}{a}\sin\theta\cos\theta\left\{-\frac{4\pi}{(2N_2+N_1)\,e^{\xi_0}}\left[-\sinh\xi_0\frac{dSek_1(\xi)}{d\xi} + \cosh\xi_0\,Sek_1(\xi)\right]\right. \\
&\quad \left. + \frac{4\pi}{(2M_2-M_1)\,e^{\xi_0}}\left[-\cosh\xi_0\frac{dCek_1(\xi)}{d\xi} + \sinh\xi_0\,Cek_1(\xi)\right]\right\}, \\
&\qquad\qquad \xi = \xi_0.
\end{aligned}\right\} \qquad (7\cdot1)$$

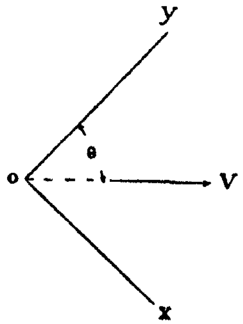


Fig. 3

Assuming, as a sufficiently close approximation, that

$$Cek_1(\xi) = K_1(k\cosh\xi), \quad Sek_1(\xi) = \tanh\xi\,K_1(k\cosh\xi), \qquad (7\cdot2)$$

numerical results for two values of a/b are given. The corresponding integrals have been evaluated by numerical integration:

(1) $a/b = 5\cdot08$, $\xi_0 = 0\cdot2$, $2M_2 - M_1 = 4\cdot15$, $2N_2 + N_1 = 3\cdot6$, $R = 2a/c \approx 2$

$$C_D = 0\cdot12\cos^2\theta + 2\cdot2\sin^2\theta,$$

$$C_L = 1\cdot0\sin 2\theta.$$

The maximum value of C_L occurs at $\theta = 45°$, and the maximum value of C_L/C_D at $\theta = 13°$ at which angle $C_L/C_D \approxeq 2$.

(2) $a/b = 1{\cdot}25, \quad \xi_0 = 1{\cdot}1, \quad 2M_2 - M_1 \approxeq 2N_2 + N_1 \approxeq 2{\cdot}14, \quad R = 3{\cdot}3.$

$$C_D = 6{\cdot}3\cos^2\theta + 9{\cdot}7\sin^2\theta,$$

$$C_L = 1{\cdot}7\sin 2\theta.$$

The maximum value of C_L is again at 45°, and the maximum value of C_L/C_D occurs at about 39°, at which angle $C_L/C_D = 0{\cdot}218$.

In fig. 4 are curves for C_L, C_D and C_L/C_D as functions of θ.

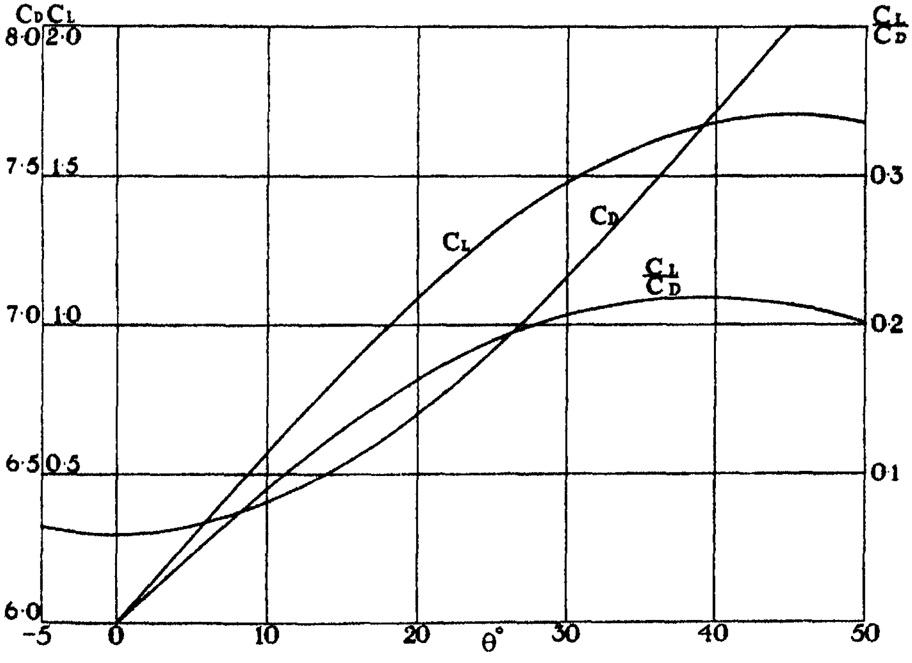


FIG. 4

For $\theta = 0$ the above values of drag can be compared with the corresponding results for an elliptic cylinder obtained by Bairstow, Cave and Lang (1923), namely

$$R = \frac{4\pi\mu V}{\dfrac{a}{a+b} - \gamma - lg\dfrac{V(a+b)}{8\nu}}, \tag{7·3}$$

where $a > b$, and the flow is along a.

The values of C_D for $\theta = 0$ are, according to (7·4),

$$C_D = 5{\cdot}84, \quad a/b \cong 5,$$

$$C_D = 7{\cdot}86, \quad a/b = 1{\cdot}25.$$

The agreement is satisfactory only in the latter case. This arises because the expansions of $Cek_1(\xi)$ and $Sek_1(\xi)$ are more rapidly convergent for large values of ξ. Now, $\xi_0 = 0{\cdot}2$ for $a/b \cong 5$ and $\xi_0 = 1{\cdot}1$ for $a/b = \frac{5}{4}$, and the one term approximations for $Sek_1(\xi)$ and $Cek_1(\xi)$ prove to be satisfactory only for the latter case.

Also, the neglecting of higher harmonics, namely ce_2, se_2, etc., is hardly justified in the case when $a/b = 5$.

In the case when $a/b = 1{\cdot}25$, on the other hand, the omitted terms are of lower order of magnitude, and the obtained results can be considered as valid to a first approximation.

I wish to thank Professor L. Bairstow, F.R.S., and Dr. W. G. Beckley for valuable help and many useful suggestions.

SUMMARY

The equations of motion of a viscous fluid round an inclined elliptic cylinder have been solved by Oseen's approximation, the solution being expressed in Mathieu functions. The complete stream function for no slipping on the surface of the cylinder has been found and formulae for lift and drag at very small Reynolds numbers < 4 have been obtained. It remains for the future to extend the computations to larger values of Reynolds number.

Incidentally, an expression is found for Green's function for a plane external to an ellipse.

For an ellipse with the ratio of major and minor axes of 1·25 and Reynolds number 3·3 the calculated values of the lift and drag coefficients and the ratio of lift to drag are shown in § 7. A maximum lift coefficient of about 1·7 is reached at an angle of incidence of 45°.

The drag increases continuously with increase of angle of incidence and the ratio of lift to drag has a maximum of about 0·22 at an incidence of 39°. Whilst no experimental results exist for such small Reynolds numbers the general type of curve is that for aerofoils.

REFERENCES

Bairstow, L., Cave, B. M. and Lang, E. D. 1923 *Philos. Trans.* A, **223**, 383–432.
Berry, A. and Swain, L. M. 1923 *Proc. Roy. Soc.* A, **102**, 766.
Courant, R. and Hilbert, D. 1931 "Methoden der Math. Physik", p. 330. Berlin: Julius Springer.
Faxén, H. 1927 *Nova Acta Soc. Sci. upsal.* Volumen Extraordinarem.
Goldstein, S. 1927 *Trans. Camb. Phil. Soc.* **23**, 303–36.
Gray, A. and Mathews, G. B. 1922 "Bessel Functions." Macmillan and Co.
Harrison, W. J. 1924 *Trans. Camb. Phil. Soc.* **23**, 71–88.
Lamb, H. 1916 "Hydrodynamics," p. 605. Camb. Univ. Press.
Piercy, N. A. V. and Winny, H. F. 1933 *Proc. Roy. Soc.* A, **140**, 543.
Ramsay, A. S. 1920 "Hydrodynamics," p. 111. G. Bell and Sons.
Southwell, R. V. and Squire, H. B. 1933 *Philos. Trans.* A, **232**, 27–64.
Whittaker, E. T. and Watson, G. N. 1927 "Modern Analysis." Camb. Univ. Press.

Reversible Oxidation of Chlorophyll

BY E. RABINOWITCH AND J. WEISS

From the Sir William Ramsay Laboratories of Inorganic and Physical Chemistry, University College, London, W.C. 1

(*Communicated by F. G. Donnan, F.R.S.—Received* 29 *May* 1937)

In studying the photochemical behaviour of chlorophyll solutions we found some simple reactions of this substance which have apparently escaped attention. We observed that chlorophyll is transformed by certain oxidizing agents, e.g. ferric chloride, into a yellow product from which it can be easily regained, e.g. by the action of ferrous chloride. The equilibrium between chlorophyll, its yellow transformation product, ferric and ferrous iron is shifted by light towards the yellow form.

In this paper we present the results obtained so far in the study of the *thermal* reaction. The influence of illumination was dealt with in a preliminary paper (Rabinowitch and Weiss 1936) and will be described in detail on a later occasion.

1—Extinction Curves of Chlorophyll *a*, Chlorophyll *b*, and Ethyl Chlorophyllide

We have measured the extinction curves of chlorophyll *a*, chlorophyll *b*, and of the mixed ethyl chlorophyllides *a* and *b*. These three preparations were kindly supplied to us by Professor A. Stoll (Basle), to whom our sincere thanks are due.

The arrangement consisted of a 1000 c.p. pointolite lamp, a condenser lens, a Hilger-Müller double quartz monochromator, and a selenium "Sperrschicht" photocell (Electrocell) connected to a Zernicke galvanometer (Zc., ~ 10^{-10} amp./mm.). Two plane-parallel glass cells (1 cm. thick) were fixed on a sledge and placed alternatively between the exit slit of the monochromator and the photocell. One cell was filled with chlorophyll solution, and the other with pure solvent. I and I_0 values were determined for wave-lengths between 3500 and 7000 A, in steps from 50 to 50 A, with slits adjusted so as to isolate bands 10–20 A wide from the continuous spectrum of the lamp. In order to control the steadiness of the lamp, two I_0 readings were taken for each wave-length, with an I reading between them.

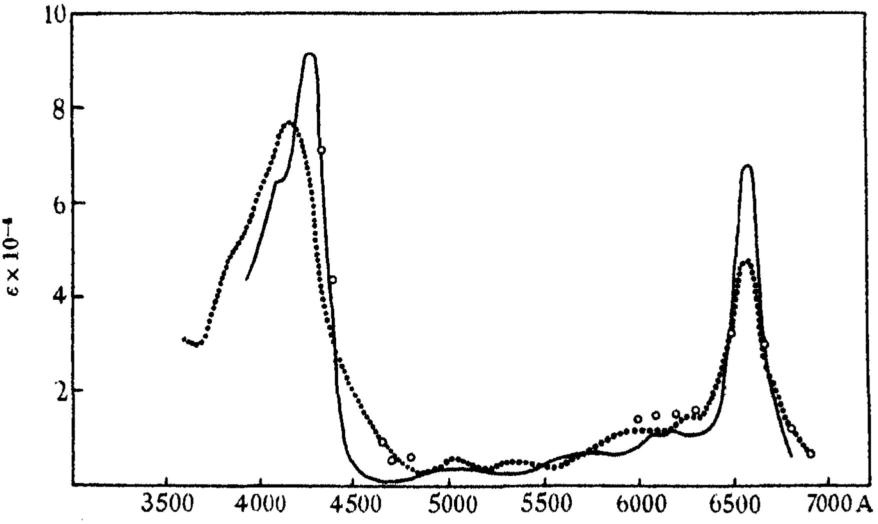


Fig. 1—Extinction curve of chlorophyll *a*.
... This paper (meth. alcohol). — Zscheile (ether).
○ Ghosh and Sen Gupta (acetone).

The samples were weighed and dissolved in methyl alcohol; they were used for the measurements not later than a few hours after dissolution. Figs. 1 and 2 show the molecular extinction coefficients ϵ ($\epsilon cd = \log_{10} I_0/I$) of the two chlorophyll components. The full lines are the curves given by Zscheile (1936). In the case of chlorophyll *a*, the two curves are roughly in agreement, although Zscheile's bands are sharper and—especially the red one—higher than ours [ϵ(6600 A) = $4{\cdot}6 \times 10^4$ according to our measurements, against $6{\cdot}8 \times 10^4$ according to Zscheile]. The position of the violet maximum is also somewhat different (4200 against 4275 A).

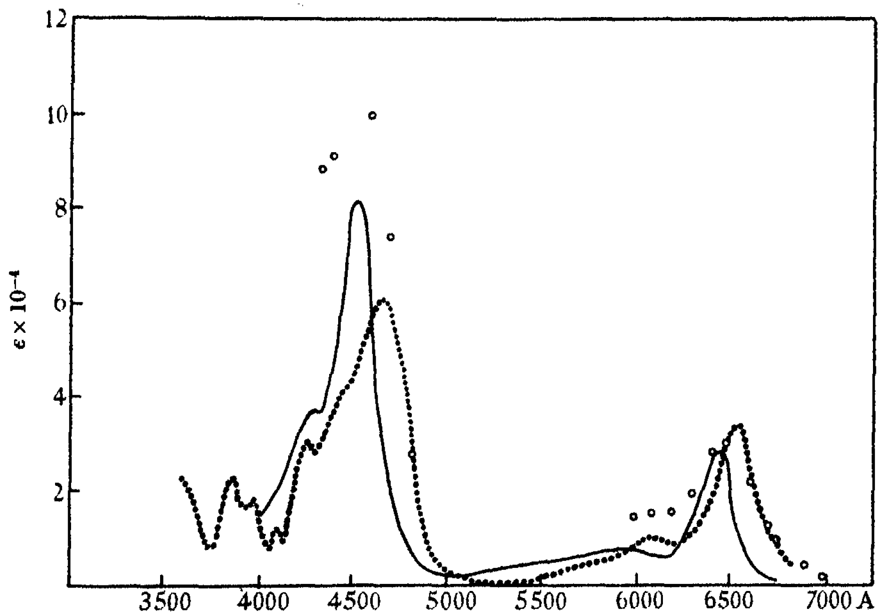


FIG. 2—Extinction curve of chlorophyll *b*.
... This paper (meth. alcohol). — Zscheile (ether).
○ Ghosh and Sen Gupta (acetone).

In the case of chlorophyll *b*, the differences are greater, the maxima of the two bands being 4650 and 6550 A, according to our measurements, instead of 4525 and 6450 A, according to Zscheile. We observed that the extinction curve of chlorophyll *b* changes with time; that of a 2 weeks old solution was much nearer to Zscheile's curve than the one determined immediately after dissolution. Fig. 3 shows this for the violet-blue band.

These discrepancies illustrate the difficulty of using extinction coefficients for quantitative analysis of *a*+*b* chlorophyll mixtures, as suggested by Ghosh and Sen Gupta (1931) and Zscheile (1936). The coefficients given by Ghosh and Sen Gupta cannot be molecular extinction coefficients because of their order of magnitude, which is much too small. If they are multiplied by 32 (so as to make the "red" maxima coincide with ours), the values marked by crosses in figs. 1 and 2 are obtained.

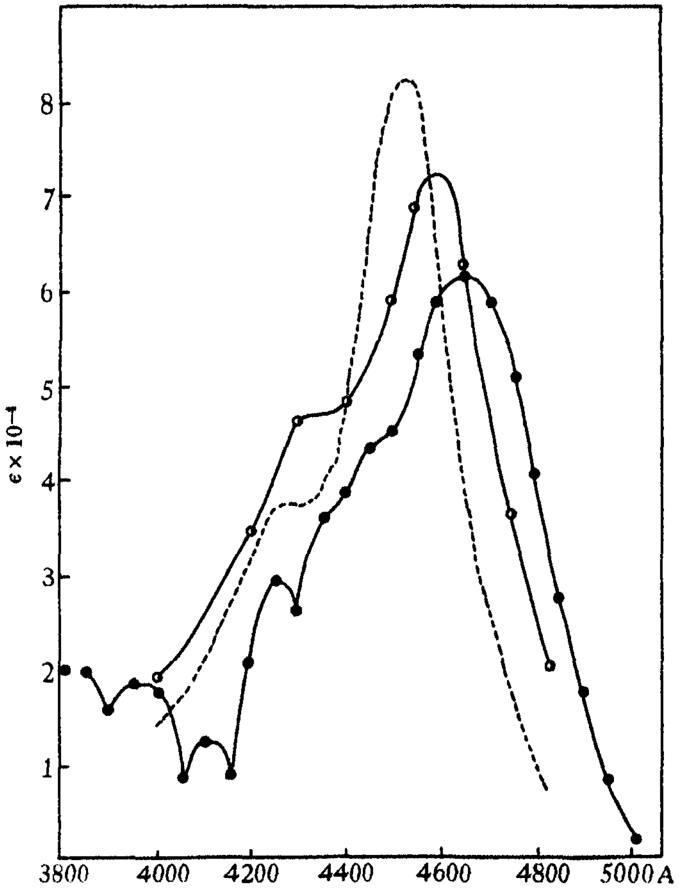


FIG. 3—Violet band of chlorophyll *b*.

● Fresh solution }
◑ 14 days old solution } in CH_3OH. - - - - Zscheile's curve (in ether).

Zscheile's curves refer to ether solutions, ours to solutions in methyl alcohol, which may account for some differences (Hubert 1935). Ghosh and Sen Gupta used acetone as solvent.

Fig. 4 shows the extinction curve of the "natural" mixture of the ethyl chlorophyllides *a* + *b*. The two maxima in violet and blue (4200 A, probably due to *a*, and 4650 A probably due to *b*), as well as the "red" maximum (6600 A) are the same as in chlorophyll. The violet band is in this case scarcely higher than the red one. This is due mainly to the coincidence of the red bands of *a* and *b*, whereas the violet bands of the two components are almost completely separated. The absolute value of the "red" maximum is, however, also higher than in chlorophyll ($\epsilon_{max} = 5 \cdot 5 \times 10^5$ for ethyl chlorophyllide $a+b$, as against $4 \cdot 5 \times 10^5$ for chlorophyll *a*, and $3 \cdot 5 \times 10^5$ for chlorophyll *b*). Stern and Wanderlein (1936) observed that the substitution of methyl for phytyl causes a still stronger increase in the intensity of the red band: the maximum extinction coefficient of methyl chlorophyllide *a* (in dioxan) in the red is 10×10^4, at 6600 A (dotted curve in fig. 4).

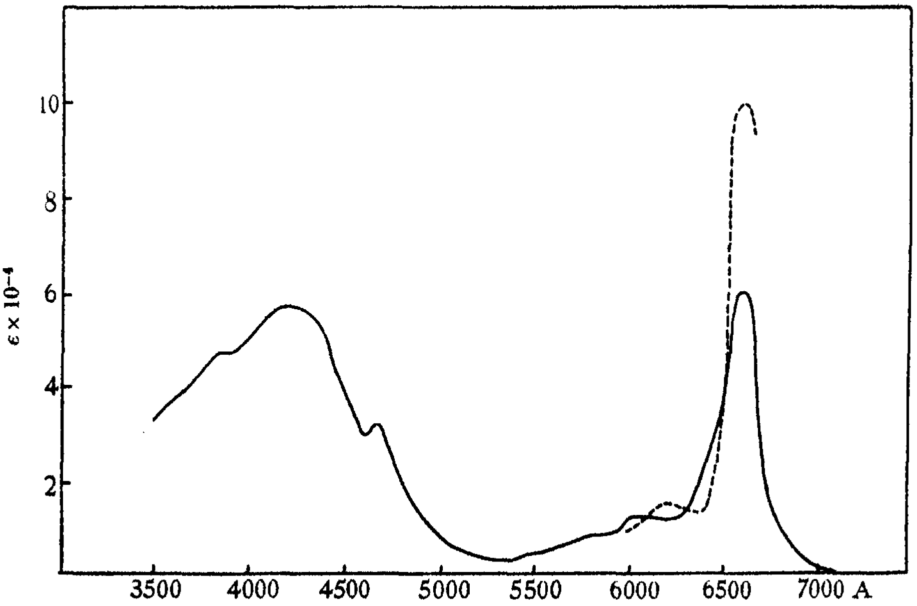


Fig. 4—Extinction curve of ethyl chlorophyllide $a+b$.

— This paper (in CH_3OH). - - - - Stern (methyl-chlorophyllide *a* in dioxan).

2—The Reaction of Chlorophyll with Ferric Chloride

All three preparations—chlorophyll *a*, chlorophyll *b*, and ethyl chlorophyllide $a+b$—immediately change their green colour into yellow on the addition of $FeCl_3$ or other ferric salt. Most of the following experiments were done in methyl alcohol solution, but the same reaction was also observed in ethyl alcohol and in acetone.

(*a*) *Extinction Curve*—The yellow colour of the reaction product suggests the disappearance of the red absorption band. This is confirmed by absorption photographs taken immediately after the reaction. A few minutes later, however, the green colour begins to reappear slowly—a process which will be

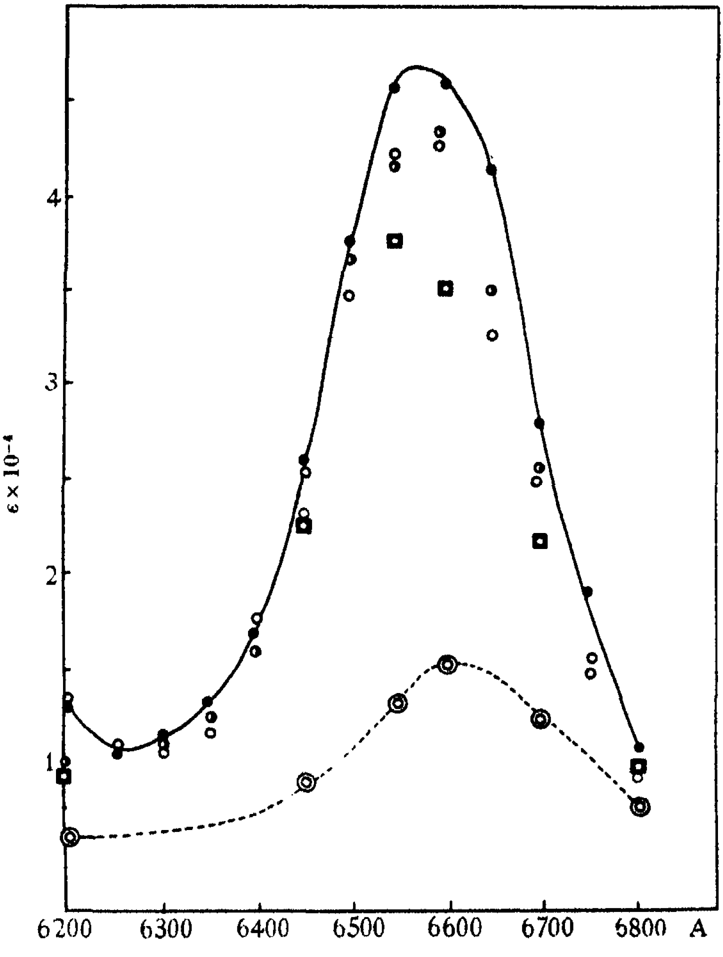


Fig. 5—Red band of chlorophyll *a*.

● Original.
◑ Restored by NaCl.
◎ Oxidized by $FeCl_3$.
◘ Restored by staying.
○ Restored by $FeCl_2$.

discussed in § 5. This "self-regeneration" of the green pigment makes the exact photo-electric determination of the extinction curve of the yellow reaction product difficult, a full set of measurements requiring about 2 hr. Fig. 5 gives a rough extinction curve of chlorophyll *a* in the red after reaction with $FeCl_3$, measured in the course of the first 5 min. after the reaction. It shows already an indication of the reappearing red band.

The violet band of chlorophyll *a* is not essentially affected by reaction (fig. 6) below 4100 A; the extinction coefficients in the interval 4100–4500 A are, however, lowered, and the maximum shifted from 4200 to 4100 A.

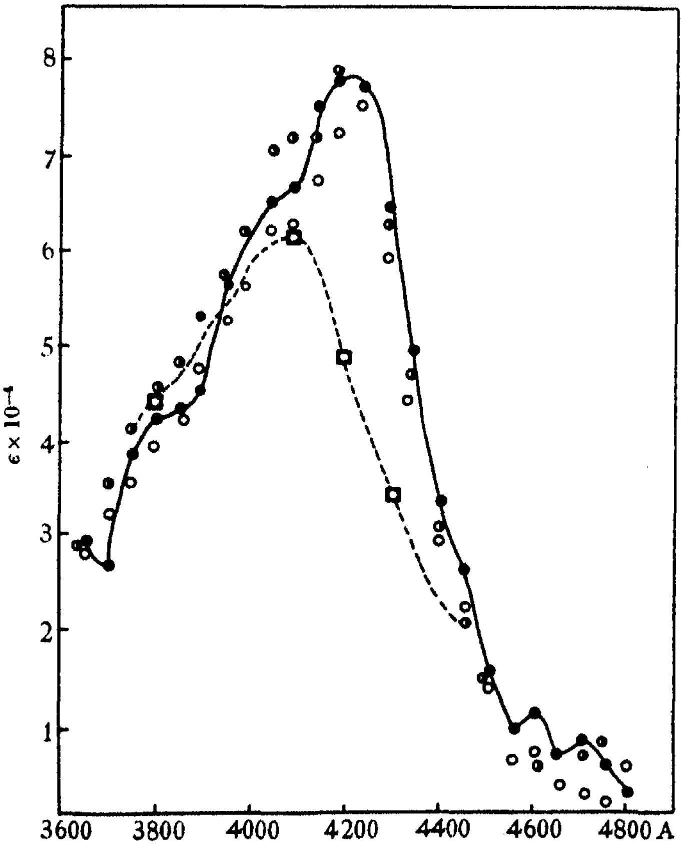


Fig. 6—Violet band of chlorophyll *a*.

● Original. ○ Restored by $FeCl_3$.
◘ Oxidized by $FeCl_3$. ◑ Restored by NaCl.

Figs. 7 and 8 show the extinction curves of chlorophyll *b* immediately after reaction with $FeCl_3$. The red band (fig. 7) has nearly vanished. The violet band (fig. 8) is more strongly affected than in the case of chlorophyll *a*; the maximum at 4650 A is lower, and a new maximum has arisen at 4300 A. (Only a slight indication of this maximum is found on the extinction curve of pure chlorophyll *b*.) This is the beginning of a *slow* reaction (obviously different from the one which causes the disappearance of the red band). Curve *C* in fig. 8 shows the appearance of the violet band 4 hr. later: only traces are left of the original maximum at 4650 A, the new maximum at 4300 A having grown at its expense. The violet band of the so transformed chlorophyll *b* lies near to that of chlorophyll *a* ($\lambda_{\max} = 4300$ A against 4200 A in *a*). The extinction curve of ethyl chlorophyllide $a+b$ after reaction with $FeCl_3$ shows that the two components behave in the same way as in the case

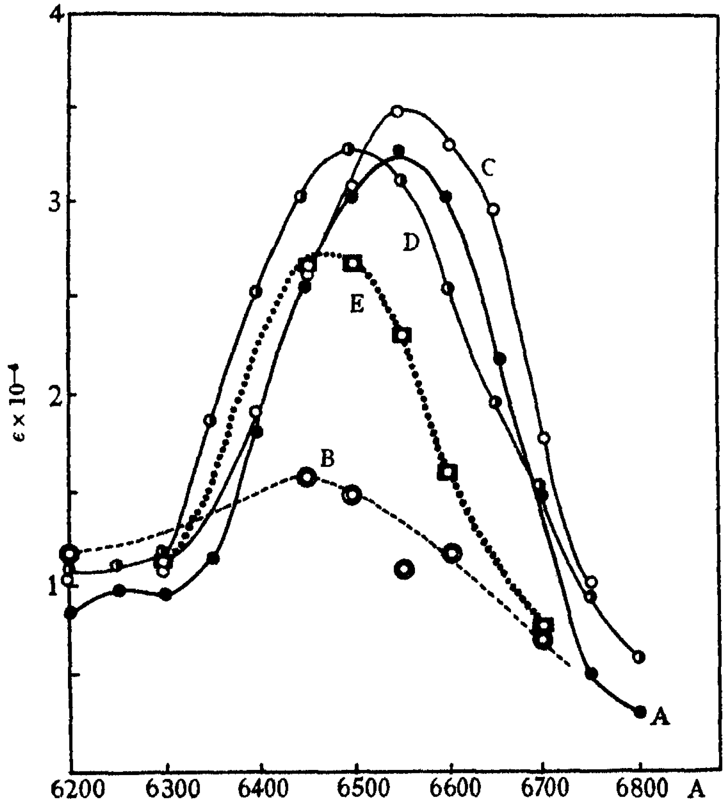


Fig. 7—Red band of chlorophyll *b*.

● Original.
◎ Oxidized by $FeCl_3$.
○ Restored by $FeCl_3$.
◑ Restored by NaCl.
▣ Restored by staying.

of chlorophyll itself. The red band disappears, the hump of the violet band at 4650 A—obviously due to the *b* component—also vanishes, and the maximum at 4200 A is enhanced by superposition of the violet band of the transformed *b* component on that of the *a* component.

These experiments show that $FeCl_3$ quickly attacks the "red" (and, much less strongly, also the violet) chromophoric group of both chlorophyll components and causes an additional slow change in the "violet" chromophoric group of the *b* component.

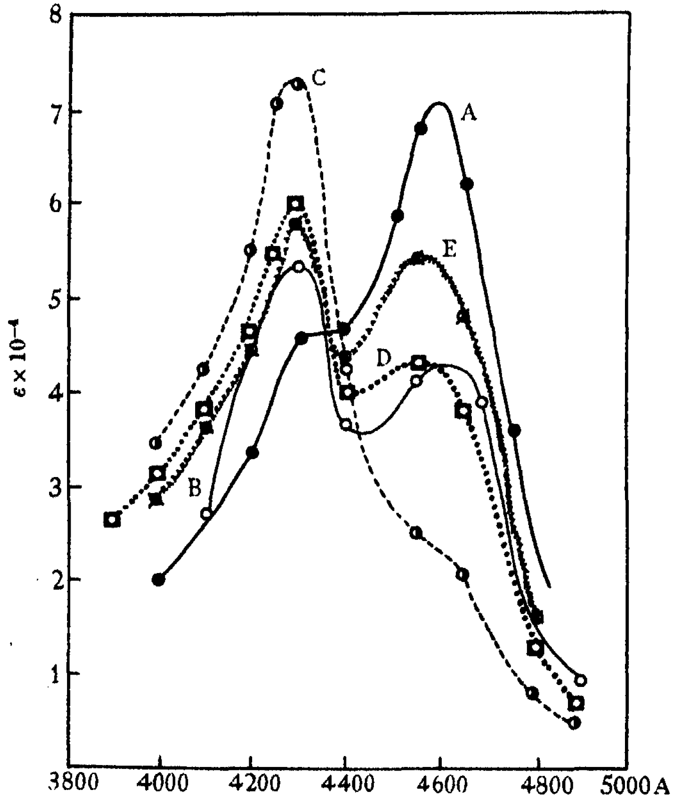


FIG. 8—Violet band of chlorophyll *b*.

A ● Original (14 days old solution in CH_3OH).
B ○ Same + $FeCl_3$ at once.
C ◐ Same, 4 hr. later.
D ▣ Same, 23 hr. later.
E ⊠ Same, 72 hr. later.

(*b*) *Fluorescence*—With the disappearance of the red absorption band, the red fluorescence of chlorophyll also vanishes. It is known that it can be excited by absorption both in the red and the violet. The yellow reaction products of chlorophyll and $FeCl_3$ still strongly absorb in the violet; fluorescence is, however, nearly absent.

The intensity of chlorophyll fluorescence can easily be measured by using complementary filters, as shown in fig. 9. With Wratten No. 47 (violet-blue) + 1 cm. saturated $CuSO_4$ solution in F_1, and Wratten No. 25 (red) in

F_2, *no light is perceptible even by looking straight into the crater of the arc*—unless red fluorescence is emitted from A. By placing the red filter, with the photocell PC behind it, in F_2', at right angles to F_1, the fluorescent light can be measured without any disturbance by scattered light from the arc.

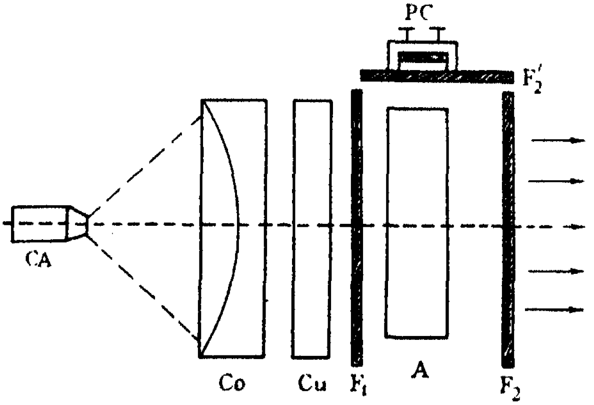


FIG. 9—*CA*, carbon arc; *Co*, condenser lens; F_1, Filter Wratten No. 47; F_2 (or F_2') Filter Wratten No. 25; *Cu*, $CuSO_4$ solution; *A*, rectangular cell; *PC*, photocell.

We found in this way, in chlorophyll + $FeCl_3$ solutions immediately after mixing, a fluorescence intensity of about 20–30 % of that of the original chlorophyll. This residual fluorescence is probably due to a certain proportion of unchanged chlorophyll being still present in the solution. More exact measurements are impossible because the yellow products are extremely sensitive to violet light (see § 3).

(*c*) *The Nature of the Reaction of Chlorophyll with Ferric Iron*—Three possible explanations of the reaction of chlorophyll with ferric chloride are:

(1) Formation of some substance containing chlorophyll and ferric iron.

(2) Isomerization of chlorophyll catalyzed by $FeCl_3$.

(3) Oxidation of chlorophyll and reduction of ferric to ferrous chloride.

The *first* assumption is improbable because none of the known compounds of chlorophyll with metals (i.e. products in which magnesium is exchanged for another metallic element) can be obtained with such ease; neither can the original pigment be regenerated directly from these substitution products (as this is possible with the substance obtained by reaction with $FeCl_3$, see below). The *second* assumption is analogous to the one suggested recently by Fischer (1936) as an explanation of the so-called "phase-test reaction", which in the case of chlorophyll *a* and its derivatives consists in a temporary appearance of a yellow coloration ("yellow phase") in the reaction of chlorophyll *a* with concentrated alkali. Fischer assumed that the yellow colour is due to enolization in the isocyclic ring. Stern (1936) and collaborators concluded, however, from extensive measurements of absorption spectra of chlorophyll derivatives, that the existence of a red absorption band is characteristic of the heterocyclic ring system (so-called di-hydroporphine system) underlying the structure of the chlorophyll molecule. Its presence is not affected by substitutions in the side-chains or in the isocyclic ring, nor by a disrupture of this ring; the red band is, however, reduced to about half its original height by removal of the complexly bound Mg, and disappears practically by dehydrogenation of the di-hydropyrrol ring. It is therefore not improbable that the "yellow phase" differs from the original chlorophyll more than by an enolization in the isocyclic ring. It is worth mentioning in this connexion that the reaction with ferric salts is shown by "aged" ("allomerized") chlorophyll solutions, which have completely lost the capacity of giving the "phase test", in the same way as by freshly prepared ones. The *third* hypothesis—which we thought to be the most probable—we tried to confirm directly by a sensitive reaction of the Fe^{II} formed by reduction. A positive result was obtained by a spot test with $\alpha\alpha'$-dipyridyl, which forms a red complex with Fe^{II}. If one drop of $FeCl_3$, one of chlorophyll and one of $\alpha\alpha'$-dipyridyl (all dissolved in methyl alcohol) are brought together on filter paper, a red spot is formed, which is absent if one of the three components has been omitted. Another argument in favour of the oxidation hypothesis is the fact that the same reaction can be produced also by other oxidizing ions, e.g. Ce^{4+} (see § 6).

The first stage in the oxidation of chlorophyll by $FeCl_3$ is in all probability the transfer of an electron from the chlorophyll to the ferric ion:

$$(1)\quad \mathrm{Chl} + \mathrm{Fe}^{3+} \rightarrow \mathrm{Chl}^{+} + \mathrm{Fe}^{2+}.$$

The cathion Chl^+ may be stable enough to persist as such in solution, subject to an equilibrium with anions present, e.g.

$$(2)\quad \mathrm{Chl}^{+} + \mathrm{Cl}^{-} \rightleftharpoons \mathrm{ChlCl},$$

or else it may dissociate, liberating an H^+-ion:

$$(3)\quad \mathrm{Chl}^{+} \rightarrow \mathrm{H}^{+} + \text{monodehydro-chlorophyll}.$$

It is also possible that the oxidation goes one step farther, leading to Chl^{2+} (or to the corresponding neutral molecule $ChlCl_2$), or else to the didehydro-chlorophyll. The reversibility of the oxidation—which will be demonstrated in the next section—can, however, be taken as an indication of it not going beyond the unstable stage—that of the radical-like odd molecule Chl^+ (or mono-dehydrochlorophyll). We shall provisionally call the "reversible" oxidation product the "oxy-chlorophyll".

We cannot say which part of the molecule is affected by the reaction. It is shown both by chlorophyll and by ethyl chlorophyllide, but no analogous change occurs with the pheophorbide. Neither a change in colour, nor a reduction of Fe^{III} can be observed in solutions in which the magnesium has been removed from the chlorophyll molecule and the latter converted into the pheophorbide by an acid. We made some experiments in order to see if positive chlorophyll ions are actually formed by reaction with $FeCl_3$, as according to (1). The mixture was placed at the bottom of a U-shaped tube; pure $FeCl_3$ solution was filled on top of it on both sides. A potential of 220 V was applied between a Pt cathode and an Ag anode. On the cathodic side a movement of the boundary between green and yellow solution was observed. After a certain time a green layer was formed half-way to the cathode, separated by a yellow zone from the green solution at the bottom. A possible explanation of these effects is that colourless (or yellow) Chl^+ ions move to the cathode and are converted back into Chl molecules by reaction with Fe^{2+} diffusing from the cathode (where it has been formed by reduction of Fe^{3+}).

These experiments are preliminary; they indicate, however, the formation of some kind of positive chlorophyll-ions by reaction with Fe^{III}.

Reaction (1) goes very quickly (in a few seconds), but not instantaneously. If solutions of Chl and $FeCl_3$ are cooled to $-80°$ C. before mixing, the change in colour is visibly delayed. Another observation showing that a certain energy is required for reaction (1) is that it is accelerated by light. This follows from experiments on the shifting of the oxidation-reduction equilibrium (15) by illumination, which were mentioned in our preliminary note (1936), and will be discussed in full on a later occasion.

3—REGENERATION OF CHLOROPHYLL BY FERROUS CHLORIDE

The yellow solution of chlorophyll oxidized by $FeCl_3$ regains its original green colour upon addition of $FeCl_2$. This reaction occurs instantaneously even at $-80°$ C. If $FeCl_2$ is added first, no change in colour at all occurs by subsequent addition of $FeCl_3$. The extinction curve of the regenerated chlorophyll *a*, is, if the regeneration has been carried out immediately after the oxidation, identical with that of the original substance, both in the violet and in the red (figs. 5 and 6). If $FeCl_2$ is added, not immediately after reaction with $FeCl_3$ but some minutes later, only a partial restoration of the red band is observed. In the case of chlorophyll *b*, the red band is restored without essential alteration (fig. 7). The differences between the curves *A*, *C*, *D*, and *E* in fig. 7 are probably due to different ages of the solution (see p. 253). The violet band remains at first in the shifted position shown by fig. 8, curve *C*.

In the case of ethyl chlorophyllide $a+b$, the hump at 4600 A remains absent and the maximum at 4100 A enhanced after the restoration of the red band by $FeCl_2$, showing that the two components of the ethyl chlorophyllide behave in the same way as in the case of chlorophyll itself.

The reaction with $FeCl_2$ presents itself, in the case of chlorophyll *a*, as a reversal of the oxidation process (1), i.e. if we assume (1) as an oxidation process:

$$(4)\quad Chl^{+} + Fe^{2+} \rightarrow Chl + Fe^{3+}.$$

In the case of chlorophyll *b*, on this quick reversible process there is superimposed a slow one, affecting the violet chromophoric group—a group which in the *b* component must be different from the corresponding group of the *a* component.

The regeneration of chlorophyll by Fe^{II} is accompanied by a restoration of fluorescence. If the reduction has been carried out immediately after oxidation, the original fluorescence intensity is restored. If, however, some minutes elapse between oxidation and reduction, the fluorescence is weaker than before, although the green colour reappears even if the solution has been kept in the oxidized state for several hours.

The regeneration of the green pigment by Fe^{II} becomes impossible after the oxidized solution has been exposed to light for a short time (violet and blue light being alone active).

An irreversible change also occurs on addition of water to the oxidized solution. The first few drops of water cause sometimes a spurious reappearance of green colour; by addition of more (say 20 %) water the colour changes to a straw yellow, which is not affected any more by addition of $FeCl_2$. Even if $FeCl_2$ was added before H_2O, the subsequent addition of water causes a slow disappearance of the green colour.

The assumption that the action of Fe^{II} is to reverse the reaction due to Fe^{III} and not merely to produce a *new* substance with an absorption spectrum very much like that of chlorophyll itself, is corroborated by the observation that the whole cycle can be repeated. By adding first $FeCl_3$, then $FeCl_2$, to a chlorophyll solution, evaporating the mixture, extracting the iron salts with water and dissolving the residue in methyl alcohol, a green solution is obtained which can again be bleached by $FeCl_3$ and restored by $FeCl_2$.

4—THE SALT EFFECT

The yellow oxidized solution of chlorophyll regains its green colour also by addition of salts soluble in methyl alcohol—e.g. $CaCl_2$, NaCl, Li_2CO_3, and KI. Figs. 5 and 6 show that in the case of chlorophyll *a* the extinction curve of the green substance obtained in this way is identical with that of the original chlorophyll. Fig. 7 shows the restoration of the red band of chlorophyll *b* by NaCl. The violet band remains shifted.

We cannot yet give a complete explanation of the salt effect. The nearest explanation would be to assume that the equilibrium

$$(5) \quad \text{Chl} + \text{Fe}^{\text{III}} \rightleftharpoons \text{oxy-Chl} + \text{Fe}^{\text{II}}$$

is shifted towards the left by increasing the concentration of anions Cl^-, because of a decrease in the concentration of free Fe^{3+} ions. Only a part of the $FeCl_3$ molecules are dissociated into ions in methyl alcohol solution, and the addition of an equivalent quantity of the strongly dissociated NaCl may cause a disappearance of the greater part of the Fe^{3+} ions. A certain difficulty is, however, presented by the fact that a corresponding—although less strong—effect can also be observed when I_2 is substituted for Fe^{III} (§ 6). Another possible explanation is that the *back* reaction in (5) is affected by NaCl. It would then be necessary to postulate that the reduction occurs by way of undissociated molecules ChlCl, formed by reaction (2), which is obviously accelerated by increasing the concentration of Cl^- ions (or other anions).

5—SELF-RESTORATION OF THE GREEN PIGMENT. CHLOROPHYLL AS CATALYST

It was mentioned in § 2 that the yellow oxidized solution of chlorophyll regains its green colour after some time, even without addition of $FeCl_2$. Fig. 5 shows the restoration of the original chlorophyll *a* after it has been left standing with $FeCl_3$ in the dark for 24 hr. (the restoration does not occur in the light because of the irreversible changes mentioned in § 3). The restoration of chlorophyll *b* is shown by fig. 7. Curves *D* and *E* in fig. 8 show that by staying in the dark for several days the blue-violet band of chlorophyll *b* is also slowly restored to its original position at 4650 A. It thus appears that the change in the *b* molecule which causes the shifting of this band from 4650 to 4300 A is also a reversible process. It is remarkable that the red absorption band is more or less completely restored by standing, although its complete restoration by $FeCl_2$ becomes impossible a few minutes after oxidation. This could be considered as indicating that the "self-restored chlorophyll" is different from the original substance. The following observations—besides the identity of the extinction curve—support, however, the restoration hypothesis: It can be shown—e.g. by reaction with NH_4CNS—that the regeneration of green colour by warming, or standing of oxidized chlorophyll for 24 hr., is accompanied by a disappearance of Fe^{III}, in quantities ten or twenty times greater than the equivalent of the chlorophyll present.

The "self-regeneration" of chlorophyll is thus due to a slow reduction of Fe^{III}. The disappearance of $FeCl_3$ is also shown by the fact that by a new addition of $FeCl_3$ the self-restored chlorophyll can be bleached again.

No change in concentration of $FeCl_3$ was observed after 24 hr. in a cell containing $FeCl_3$ solution in CH_3OH without chlorophyll. It thus appears that chlorophyll acts as a catalyst for a reduction of $FeCl_3$ in the dark—most probably by oxidation of CH_3OH. When the chlorophyll oscillates between the original and the oxidized state according to (5), a small number of molecules of the oxidized substance return to the reduced state, not by reaction with Fe^{II}, but by oxidizing the solvent. This reaction is obviously slow, i.e. it requires an activation energy. The self-regeneration of oxidized chlorophyll is delayed indefinitely by cooling to $-80°$ C., and accelerated by warming.

6—OXIDATION-REDUCTION EQUILIBRIA WITH OTHER REAGENTS

A change of colour from green to yellow also occurs when *ceric salts*, *iodine* or *cupric chloride* are added to a chlorophyll solution; *hydroquinone*, *potassium iodide* and *cuprous chloride* cause the green colour to reappear. In the case of $CuCl_2$, the "self-restoration" of the green colour occurred very quickly (i.e. in the course of a few minutes). The addition of sodium chloride causes a reappearance of the green colour, also in the reaction with iodine, although the recoloration goes much more slowly than in the reaction with ferric ions.

We see the importance of these results for the chemistry of chlorophyll—and possibly for the part played by this pigment in the photo-reduction of carbon dioxide—in the demonstration of the fact that Chl yields a highly reactive oxidation product, from which it can easily be regained by reduction. This property enables chlorophyll to act as an oxidation-reduction catalyst, even in the dark. Experiments on the influence of illumination on the oxidation-reduction equilibrium show that light absorption increases the tendency of chlorophyll to be oxidized, i.e. enhances its capacity of acting as a reduction catalyst.

SUMMARY

In this paper some experiments are described showing that chlorophyll reacts with certain oxidizing agents—especially with *ferric salts*—forming a yellow product, from which it can be easily regained, for instance, by the action of *ferrous salts*.

Photoelectric measurements were made of the extinction curves of:

I. The methyl alcohol solutions of chlorophyll (both the pure chlorophylls *a* and *b*, and the "natural" mixture of the two ethyl chlorophyllides).

II. The solutions of the same three substances after reaction with $FeCl_3$.

III. The same solutions "restored" by $FeCl_2$.

Solution (II) showed the complete absence of the characteristic red absorption band of chlorophyll, whereas in solution (III) this band was restored to its original shape.

Fluorescence of chlorophyll was shown to drop to 20–25 % of its original value by addition of $FeCl_3$, and to reappear in its original strength by addition of $FeCl_2$.

These experiments are explained by the assumption of an equilibrium $\mathrm{Chl} + \mathrm{Fe}^{\mathrm{III}} \rightleftharpoons \text{oxy-Chl} + \mathrm{Fe}^{\mathrm{II}}$, where oxy-Chl is an oxidation product which may be, for instance, a positive chlorophyll-ion, or a dehydrochlorophyll.

The oxidation theory of the reaction of chlorophyll with $\mathrm{Fe}^{\mathrm{III}}$ is confirmed by a sensitive test for $\mathrm{Fe}^{\mathrm{II}}$ in a mixture of chlorophyll with $FeCl_3$. The reversible character of the reaction is verified by the possibility of repeating the cycle after separating the "restored" chlorophyll from the iron salts.

It is further shown that chlorophyll is also restored in the solutions decolorized by $FeCl_3$ by keeping them for several hours in the dark at room temperature. This is due to the disappearance of $FeCl_3$, which apparently oxidizes the solvent (CH_3OH) in the presence of chlorophyll. Chlorophyll is thus able to act as an oxidation-reduction catalyst in the dark. The reaction of chlorophyll with ferric chloride is accelerated by illumination.

Further experiments described show the instability of the "oxychlorophyll" towards light and water, the influence of salts upon the equilibrium, the special effects obtained in the violet absorption band of the *b* component, and the analogous equilibria obtained in the systems Ce^{4+}—Ce^{3+}, $\mathrm{Cu}^{\mathrm{II}}$—$\mathrm{Cu}^{\mathrm{I}}$, and I_2—I^-.

References

Fischer, H. 1936 Mikrochemie, "Molisch-Festschrift", p. 67.
Ghosh, J. C. and Sen-Gupta, S. W. 1931 *J. Ind. Chem. Soc.* **8**, 581.
Hubert, B. 1935 "The Physical State of Chlorophyll in the Living Plastid." Thesis, Leiden.
Rabinowitch, E. and Weiss, J. 1936 *Nature, Lond.*, **138**, 1098.
Stern, A. and Wenderlein. H. 1936 *Z. phys. Chem.* A, **176**, 81.
Zscheile, Jr., F. P. 1936 "Cold Spring Harbor Symposia on Quantitative Biology", **3**, 108.

On the Nature of Atmospherics—V

By R. A. Watson Watt, the late J. F. Herd and F. E. Lutkin
Radio Department, National Physical Laboratory

(*Communicated by Sir George Simpson, F.R.S.—Received* 4 *June* 1937)

[Plates 11, 12]

Despite the defects and limitations of the methods described in the first three papers of this series (Watt and Appleton 1923; Appleton, Watt and Herd 1926), which were recognized from the first and towards whose elimination continued experimental work was directed, it was considered desirable to make a preliminary survey of the problem of the wave form of atmospherics without awaiting the results of this developmental work. Extensive series of observations were therefore made in the Red Sea and Indian Ocean, in Egypt, and in the Anglo-Egyptian Sudan, with "control" observations in south-east England. In all of these series one or more of the present authors participated as observers. While no detailed discussion of the results is at present proposed, the general features of the survey are presented in Tables I–VI (*see* pages 286–291), in the form which was adopted for Table II of the second paper of this series.

The authors' own criticisms of this early technique fell into two main groups: those arising from the impossibility, at the time of the early observations, of obtaining photographic records on any economically attainable basis of adequate sampling, and those arising from the lack of sufficient knowledge as to the place of origin of the disturbances studied. Some of the criticisms made by other workers, however, appeared to be based on misconception of matters discussed in the earlier papers. It is not now necessary to discuss these criticisms at length; one of the most important points at issue was very fully discussed at the Copenhagen meeting of the Union Radio Scientifique Internationale, and this discussion resulted in the complete acceptance of the argument of the papers, that the time constant of the input circuit, if sufficiently high, did not modify the apparent duration of the field change when the initial and final field values were sensibly the same. *Ad hoc* photographic tests by the authors, on large-aerial systems of the type used in the earlier work, in which the time constant was varied over wide limits had, in the interval, confirmed the view that sensible distortion of the recorded form only occurred when the time constant fell to the abnormally low value of about four times the duration of the field change to be delineated.

The improvements in technique which have enabled us to resume discussion of the nature of atmospherics have already been outlined in a publication in which two of us were associated as authors with our colleague Mr. L. H. Bainbridge-Bell (Watt, Herd and Bainbridge-Bell 1933), and they need not, therefore, be discussed here. Mention will, when necessary, be made of outstanding problems of technique in relation to results illustrating or suggesting the need for special care in interpretation.

The present discussion is concerned with two main series of observations, in which the wave forms and apparent directions of arrival of individual atmospherics were recorded photographically at the two ends of the 560 km. base-line Slough (51° 29′ N., 0° 34′ W.)—Leuchars (56° 23′ N., 2° 52′ W.). The Slough installation for the delineation of wave form comprised a T-aerial of 25 m. geometric height, with a single-wire top 500 m. long, connected to the resistance-battery coupled amplifier described in Paper II of this series.

The output from the amplifier was arranged to produce horizontal deflexion of the cathode-ray spot and the resulting patterns were photographed on moving film which provided the time scale for opening up the patterns. Film speeds of ½ m./sec. and later 1 m./sec. were used, and these speeds, with a fine-grained film emulsion, with the new cathode ray tubes giving a very much finer spot than had hitherto been obtainable, and with a camera lens of high resolving power, enabled frequencies up to a few kilocycles per sec. to be sufficiently opened for analysis.

At Leuchars the aerial system was of smaller dimensions than that at Slough, the geometric height being 15 m. and the single-wire roof 65 m. long.

The pairs of photographs reproduced in figs. 3*a*, *b*, *c*, Plate 11, are perhaps the most directly convincing evidence that the recorded wave forms at the two stations are not sensibly influenced by the substantial differences in aerial and grid-circuit characteristics at the two places.

At both stations the photographic recording of apparent directions of arrival was carried out on identical installations of the type described by Watt, Herd and Bainbridge-Bell (1933, pp. 140–71).

The substitution of a moving-film technique for the earlier time-base system was justified by the complete sampling which it permitted, in respect both of full delineation of each individual field change and of the whole of the changes occurring during the observations, and also by the simplification which it permitted in the heavy labour of interpretation and measurement.

The operations were controlled from Slough by means of a line telephone link kindly made available by the Engineer-in-Chief of the Post Office; this line was used for the transmission of time-marking impulses so that all four films (films of wave form and direction of arrival at each end of the base-line) bore time marks from a signalling apparatus at Slough. This apparatus impressed short trains of an oscillation of 1000 c./sec. on all four oscillographs at intervals of 1 sec. (in later work at intervals of 0·2 sec.), with characteristic signals at each half-minute and minute. The certainty with which an individual atmospheric can be identified at each station, and the substantial identity of the principal features of the disturbance stream at the two stations 560 km. apart are illustrated by figs. 4*a*, *b*, *c* and *d*, Plate 12.

The first series of records to be discussed was made during August and September 1931. In this series simultaneous directional records were made at both stations, but wave-form records at Slough only. The film speed for directional recording was 2 cm./sec., for wave-form 50 cm./sec.

In the course of this first series the increased prominence of the high audio-frequencies in the forms delineated photographically led to the replacement of the resistance-battery coupled amplifier by the resistance-capacitance coupled amplifier shown in fig. 1.

The main reason for using the resistance-battery coupling had been the need for a good response characteristic for the lowest frequencies. This involved a sacrifice of the response to frequencies above about 10 kc./sec., since the capacitative shunt paths due to the necessary bulk of the coupling batteries permitted the higher frequencies to escape amplification. Now the observations already discussed have established the fact that the lowest frequencies attaining sensible amplitudes in the spectrum of the delineated "radiation" forms were over 100 c./sec. Further, the spurious effects due to the grid-circuit time constants prevented full quantitative utilization of the trace delineating "net-change" characteristics. Thus it was clear that a resistance-capacitance cascade permitting a much improved response to the higher frequencies which it was now desired to record would still give adequate low-frequency response, while the provision of amplifier time constants of the same order as those of the input grid circuit would not impair the value of the records to any greater extent than did the grid-circuit limitations alone. The frequency-response curves of the amplifiers adopted are shown for several gain levels in fig. 2.

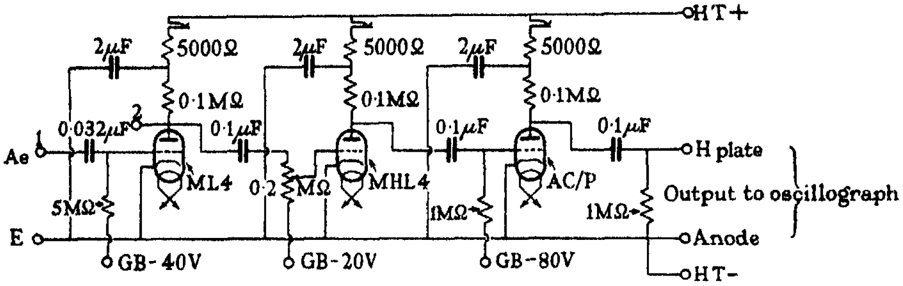


FIG. 1—Circuit diagram of amplifier. 1, 2—High-insulation low-capacity terminals.

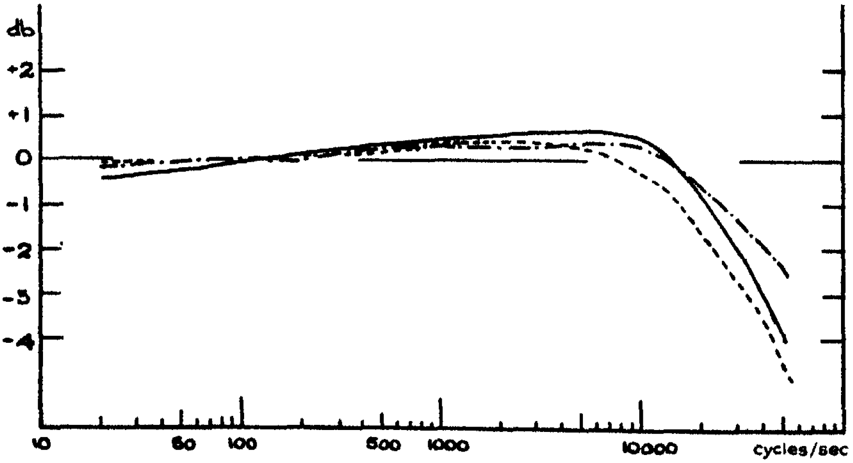


FIG. 2—Amplifier characteristics at various gain values.
—— full gain; – – – –14 db; —·—· –26 db.

The available amplification with this new amplifier was of the order of 1000 times, so that adequate amplification of atmospherics of much smaller amplitude than could be recorded with the old amplifier was obtained.

The second series of observations, in which both types of record were made at both stations, was made in the period December 1933–January 1934. The speed of the wave-form recording film was increased to 1 m./sec. during this period for better resolution of the higher frequency components. The relatively low level of atmospheric disturbance in this winter period involved the use of high gain levels in the amplifiers, so that in this series some inconvenience resulted from the recording, superposed on the atmospheric wave forms, of signals from long-wave high-power radiotelegraphic stations, and of deflexions due to the residual voltage from electric supply networks. At each station the 50 c./sec. voltage induced into the recording installation from the supply mains was approximately balanced out by a variable e.m.f. from a phase-shifting transformer, but irregularities of supply wave form and variation with the load on the mains prevented the complete maintenance of compensation. Some small proportion of the delineated forms were rejected because of the impossibility of freeing their measurement from doubts due to this superposition.

The examination of the photographic records from the earlier series revealed the fact that while the prominent general features of individual forms had been delineated correctly in the earlier eye-and-hand work, a very important relation between form and distance of origin had, on account of the recognized limitations already mentioned, escaped detection. The examination of records from sources at distances varying from approximately 200 to 4000 km. confirmed the previous conclusion that the typical composite atmospheric* consisted of a succession of comparable elements, separated by time intervals of the order of 10–50 msec. (cf. fig. 6). Each such element, however, was now found to comprise a "slow" component very closely comparable with those delineated in the earlier work, and an "oscillatory" component, containing relatively high-frequency elements, also comparable with forms previously delineated. The essential novelty of these 1931 results was that the temporal incidence of the higher frequency elements relative to the lower was controlled by the distance between the observing station and the discharge channel. In very few cases were "slow" forms recorded without traces of an accompanying "oscillatory" form, although in a considerable number of cases the "oscillatory" form was delineated without a corresponding "slow" form. In such cases the whole atmospheric seemed to consist of a succession of the simple oscillatory trains, separated by quiet periods of the kind mentioned.

---

* We do not propose to enter into any long discussion of what is meant by the term "an atmospheric". Such discussion as has been published in the interval between Papers III and IV of this series seems to us to be unnecessary and undesirable in the present state of our knowledge. We propose to apply the term "an atmospheric" indiscriminately to the relatively long-sustained group of field changes which the radiotelegraphist calls an atmospheric, and which one of us has shown to persist in some cases over a period of 1 to 5 sec. (Watt 1927, 1929) and to the shortest discrete component of the group, a component which may have a duration of a fraction of a millisecond. In any adequate discussion the context should obviate confusion, and there is as yet no sound physical basis for a closer definition.

In the case of atmospherics radiated from sources within some 200 km. of the recording station the oscillatory component formed a high-frequency embroidery on the slow form, as illustrated in the earlier work. When, however, the discharge channel was at a distance of the order of 2000 km., the high-frequency components appeared to have outrun those of lower frequency, so that the typical element then comprised an oscillatory train followed by a slow form which was completely detached from it by an interval of sensibly zero field-change, this interval being of the order of a single millisecond.

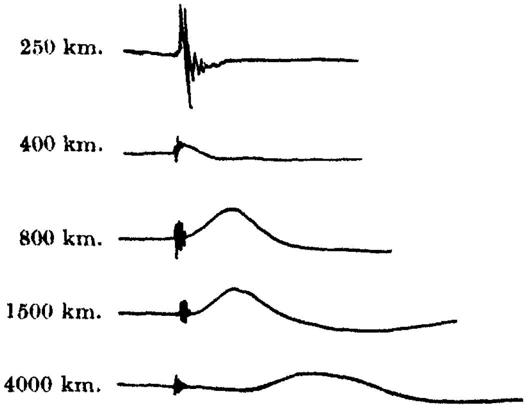


FIG. 5—Tracings of typical forms showing the detachment of the two components with increase of distance of origin.

A typical group of records, traced from enlargements of the original photographs, is reproduced as fig. 5; these show the characteristics which have just been outlined. While this first series is of historical importance, as that in which this very important relation between form and distance was first detected and studied, further improvements in technique make it desirable to give more attention in discussion to the later series than to these earliest photographic studies. It was, in fact, thought desirable to delay publication of the results obtained in the reduction work, which was carried out in 1932, until the improved methods then being developed could be applied to the closer study of this important dispersive phenomenon. The relation between form and distance was of the kind to be expected from

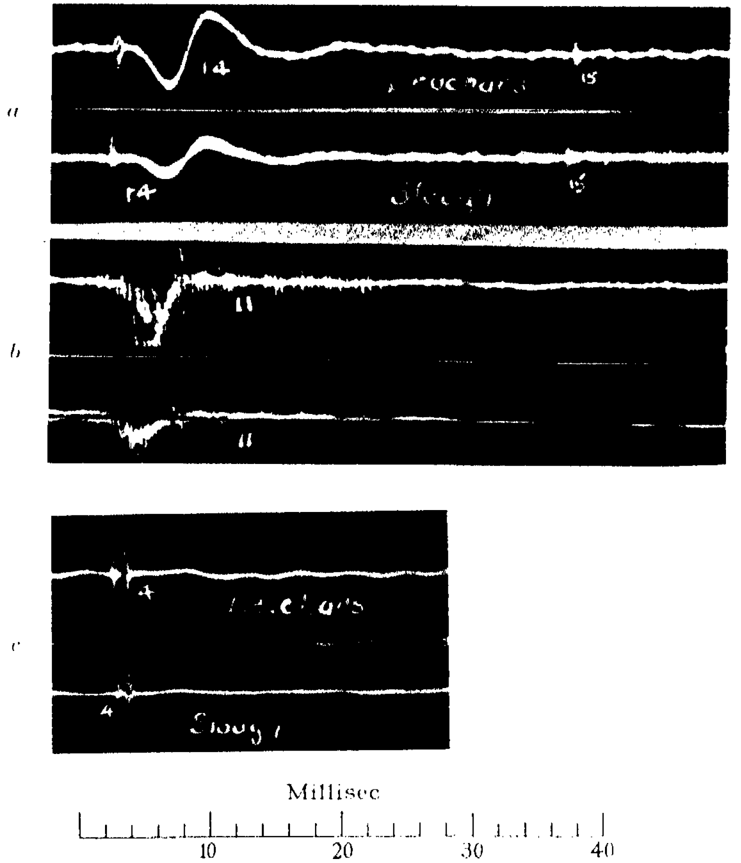
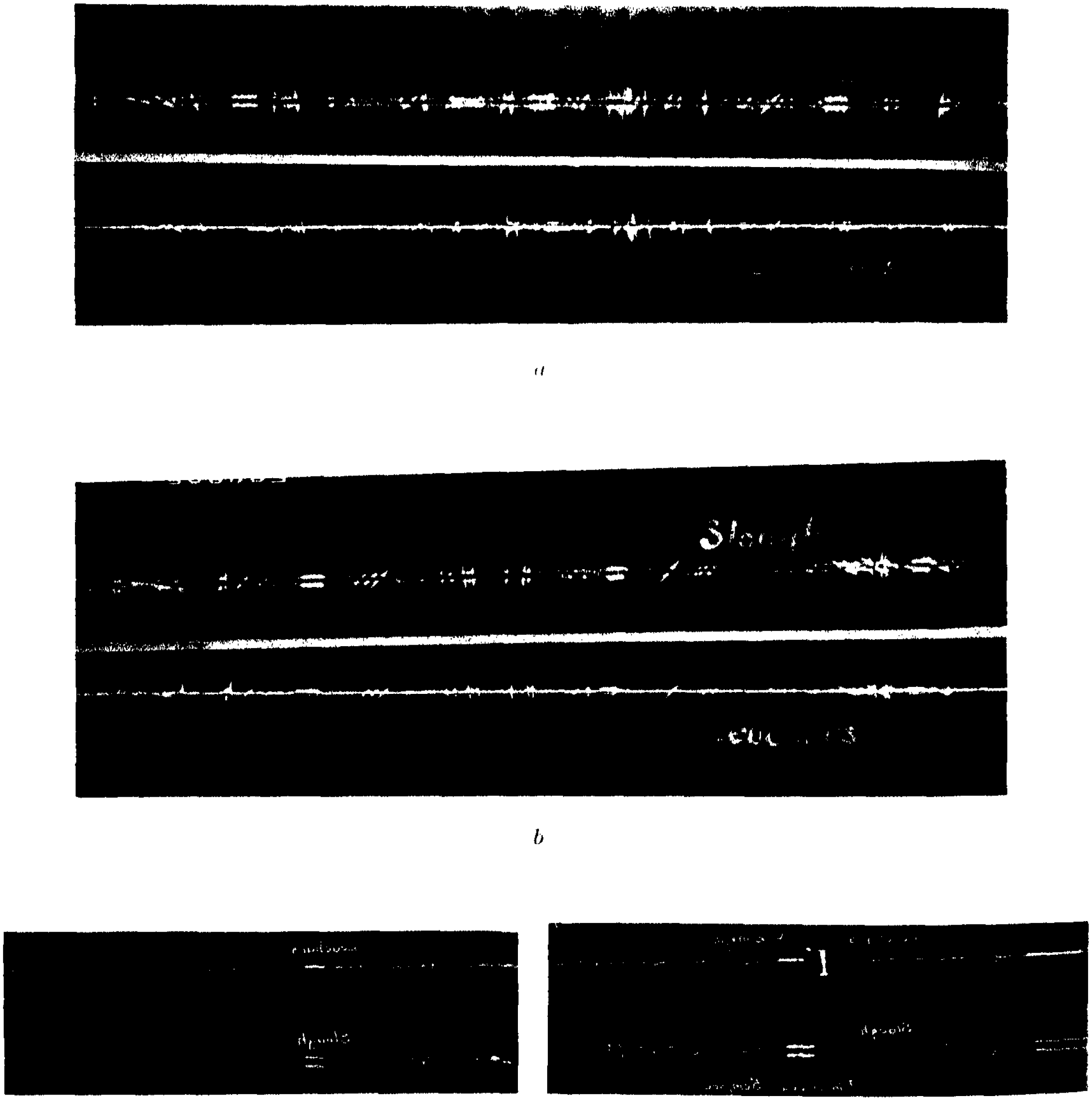


FIG. 3 Typical wave forms recorded simultaneously at Slough and Leuchars. Showing similarity of forms at the two stations.



a

b

c

d

FIG. 4. Simultaneous records of the direction of arrival of characteristic groups of atmospherics at Slough and Leuchars.

propagation in a dispersive medium, and, since it was thus likely to throw light on the still wider problem of propagation of radio waves through the atmosphere, greater temporal accuracy and greater accuracy in location of sources were very important.

In the present part of this discussion it is proposed to deal separately with the observed characteristics of the slow and oscillatory components, and then to proceed to the relations between them and to their relations with the net change of field corresponding to the "electrostatic" term.

THE SLOW COMPONENT

It was already strongly indicated in the earlier work that the division into "quasi-periodic" and "aperiodic" forms did not correspond to any profoundly fundamental phenomena, the apparently aperiodic form having most of the quantitative characteristics of the first half-cycle of the quasi-periodic form. This division has therefore been abandoned in the present discussion—although the systematic differences observed require later discussion—and the characteristics will be specified in terms of a first quasi-half-cycle, which may be followed by a second and later half-cycles of sensible or of negligible amplitude. The *duration* of the first half-cycle varied from 0·5 to 6 or 7 msec.; in the disturbances of small amplitude no great accuracy in measurement of the points of departure from and return to the base-line was attainable. The mean value of the duration of the first half-cycle in the "summer" (1931) series was 2·8 msec., in the "winter" (1933–4) series, 5·0 msec. giving an overall mean of 3·7 msec. Later reference will be made to the difference between summer and winter means.

The peak field intensity recorded on the average was, as explained in earlier papers, largely determined by the upper and lower limits of recording. The gain levels used on any particular occasion were set with regard to good sampling and economical utilization of film for that occasion, so that while in the summer runs full-scale deflexion corresponded to several volts per metre, the corresponding value for winter runs was about 0·1 V/m. The average peak field-change recorded in summer was 263 mV/m., in winter 7 mV/m., the overall mean 163 mV/m.

SIGN RELATIONS

Of 370 slow forms observed in the summer series, 296 had the first half-cycle positive, 74 negative; the 235 winter forms gave 176 positive to 59 negative. In both series, then, it will be seen that positive forms were more

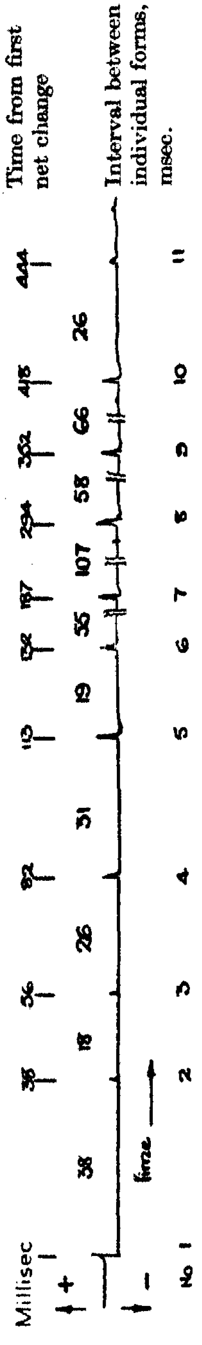


FIG. 6—Tracing showing a typical atmospheric group with eleven component discharges occurring in less than half a second.

than three times as numerous as negative, the preponderance being greater in summer. The visual series previously reported had shown a value of 1·4 for this ratio. The considerable difference in the summer photographic series may in part be referred to the fact that a frequent and typical summer phenomenon is a sequence in which a large negative is followed by a group of smaller positives, with time intervals of the 5–50 msec. order already mentioned, as shown in the specimen trace of fig. 6. In eye-and-hand recording the negative first discharge would, alike from amplitude and time of incidence, be preferentially recorded.

The mean peak value of negative first half-cycles was 1·4 times that of positives.

The oscillatory component was present in all but about 1 % of the recorded forms. It usually took the form of a train of some eight or ten oscillations of sensible amplitude, and the logarithmic decrement of a damped oscillation of approximately constant frequency most nearly representing the observed decay rate was tabulated for each form by visual matching against reference curves. An attempt to specify the sign of the initial departure was abandoned as arbitrary; changes of gain-level would clearly bring into view or suppress the one or two oscillations of small amplitude which were involved in the building-up process. The average peak amplitude of the summer series was ·340 mV/m., winter 32 mV/m., of the whole group 172 mV/m. The equivalent logarithmic decrements were for summer 0·8, winter 0·4, overall 0·5.

The frequencies which were prominent within the oscillatory component were of the order of 5–20 kc./sec., the quasi-period of the earliest arriving oscillations being shorter than that of the later parts of the oscillatory group. The assigning of a mean quasi-frequency by the counting of the number of principal stationary points in the whole duration of the oscillatory group would have produced an arbitrary relationship with amplitude, as the extreme earliest and latest, and lowest amplitude, portions of small amplitude groups would have escaped measurement. The "prominent quasi-frequency" tabulated is therefore that of the oscillations occupying the period of about 0·5 msec. centred on the peak amplitude, this peak occurring normally, but not invariably, very near the beginning of the received field change. The mean prominent quasi-frequency in summer was 5·5 kc./sec., in winter 10 kc./sec., overall 8·5 kc./sec. In the 1931 summer series the resolution was inadequate, and this may contribute to an explanation of the low summer value, but other factors will be considered later.

It will in fact be found that the steady improvement of technique has permitted the photography, since the second main series here discussed, of random samples in which a much clearer view of the frequency distribution is possible; these will be discussed in a later paper.

RELATIONS WITH DISTANCE

Passing from these general indiscriminate averages, we may now proceed to use the data as to approximate distance of origin, which forms the second important new feature of the present series. No very high precision can be claimed for the location of the place of origin of any particular atmospheric, for the following reasons. The location is effected by coil direction-finders; it is subject, therefore, to abnormal-polarization error. Only two stations are used; the fix is subject, therefore, to doubts in triangulation which would have been resolved by a third bearing. The base-line is of 560 km. only, running nearly North and South; it is too short, therefore, for accurate location of distant sources, and is ill-oriented for location of nearer sources, since English sources can never subtend any great angle relative to the base-line. The present application is not, however, one in which high precision is required, and it is not considered that any feature of the present very general review is in substantial doubt on account of errors in location. The data do, however, indicate, as has already been noted, a wide field of research into the propagation of low radio frequencies, in which more precise location by direction-finding will be required for satisfactory progress. It will become increasingly obvious, as discussion proceeds, that the irregularities in mean curves of the different parameters, and the scatter about these means, are mainly due to the relatively steep variation with frequency of the two main factors in propagation, group velocity and attenuation. The present review can hardly go further than to show how strongly marked are these relations with frequency, even in the very narrow frequency band of about 0·5–20 kc./sec. studied by the methods in use. The number of sampling periods discussed does not, for example, permit of any discussion of the diurnal variation of these propagational characteristics.

Stringent investigation of the dispersive process which leads to the separation of the atmospheric into two discrete portions cannot yet be attempted. Qualitatively the phenomena are relatively simple, but the interpretation of the quantitative data is made difficult by their incompleteness and diversity. Thus the scatter diagram of fig. 7 shows the variation, with distance from place of origin, of the time interval between the first sharply marked excursion identifiable as belonging to the oscillatory component, and the corresponding first excursion initiating the slow component. But this diagram could only be reduced to a specification of group velocities were the form of the field change at the place of origin, and the form at the receiver, completely known. The former is unknown, but is known to be far from constant. The diagram quite certainly includes points representative of many different forms at origin; the relative phases of the Fourier components in the originating field change are certainly

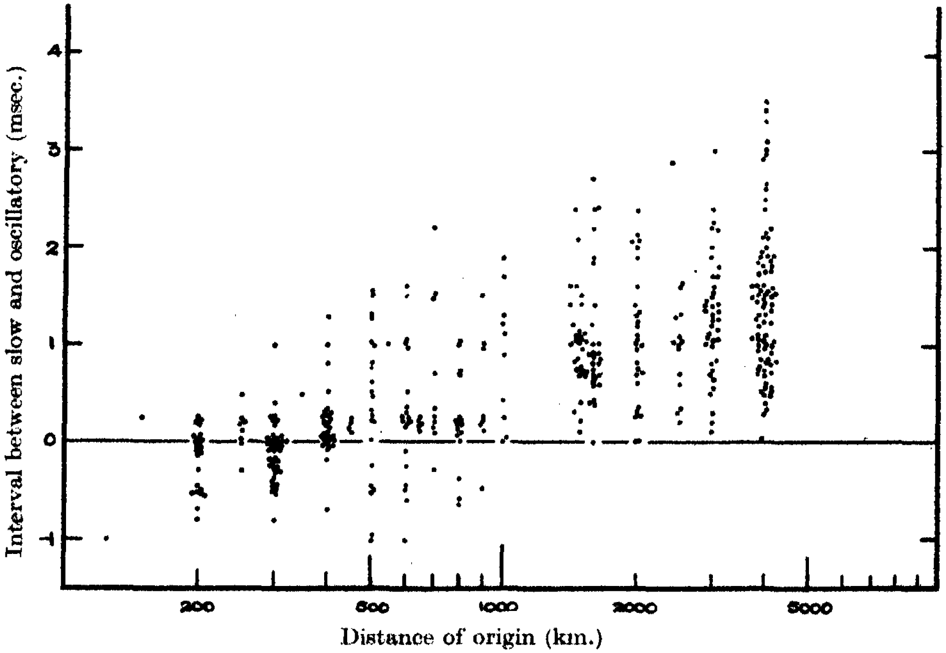


Fig. 7—Scatter diagram showing the relation between distance of origin and the interval between the slow and oscillatory components.

not constant from one discharge to another, even in the case of discharges closely grouped in space and time. The form at the observing station is known approximately, but the resolving power of the installations used in the series now under discussion was still inadequate for full specification.

The nature of these difficulties is well illustrated by one specimen group of observations. On 21 December 1933, a single run of 30 sec. duration showed nineteen atmospherics as proceeding from one small area of about 1500 miles radius centred in the Atlantic, at 45° N., 23° W. The peak amplitudes of these nineteen, as recorded at Slough, varied from 10 to 300 mV/m. One which had a peak of 150 mV/m. at Slough had a peak of 50 mV/m. at Leuchars; another of 40 mV/m. at Slough reached 40 mV/m. at Leuchars, and the ratios of peaks at the two places varied from 0·3 to 1·0. No simple law relating amplitude with predominant frequency will reconcile these values with the assumption of a constant "standard specification" at origin; correspondingly it may be concluded that the possibly simple law relating group retardation with frequency will be obscured in the data by this diversity at source.

The difficulty of interpreting fig. 7 may be stated more directly by saying that the measured interval will contain two parts, one due to the relative lateness in time of occurrence at source of the first high-frequency element, the other due to the group retardation of the low-frequency elements relative to the high-frequency element, over the trajectory between source and observer. The first part is unknown, and variable from one case to another, the second part depends on the spectrum at source. There is, moveover, a fundamental difficulty in explaining why there should be two discrete parts in the atmospheric of distant origin, with an interval of zero field change between. The separation into two parts and a gap, as opposed to a single long train with high frequencies merging gradually into low, indicates a gap in the spectrum of the atmospheric, but whether the gap is in the spectrum at source or is mainly due to selective absorption, en route, of the frequencies between about 1000 and 5000 c./sec., has yet to be explored. The hypothesis of selective absorption by the intervention of hydrogen ions would be more attractive were there any reasonable ground for postulating the existence of hydrogen ions in quantity.

It is, then, unprofitable to go further at this stage in the handling of fig. 7 than to say that the group of frequencies round 500 c./sec. appear to travel with group velocities some 10 % less than those around 5000–10,000 c./sec.

The ambiguities due to variation in spectrum are illuminated by an examination, which again can be only preliminary and rudimentary, of the relative amplitudes and frequencies at Slough and Leuchars. Over the whole of the data the peak amplitude at Leuchars, compared with the peak amplitude at Slough in the same atmospheric, varied from 0·2 to 2·0. In the specific group already quoted, this ratio varied from 0·3 to 1·0, and the plotting of this amplitude ratio against predominant frequency gives a curve, fig. 8, of remarkable steepness but reasonable smoothness. It must be inferred from this curve that any attempt to investigate attenuation in relation to frequency and distance only would be unsuccessful, and that the problem must be faced with the realization that the relative attenuations, and the closely associated relative group velocities, are dependent on some

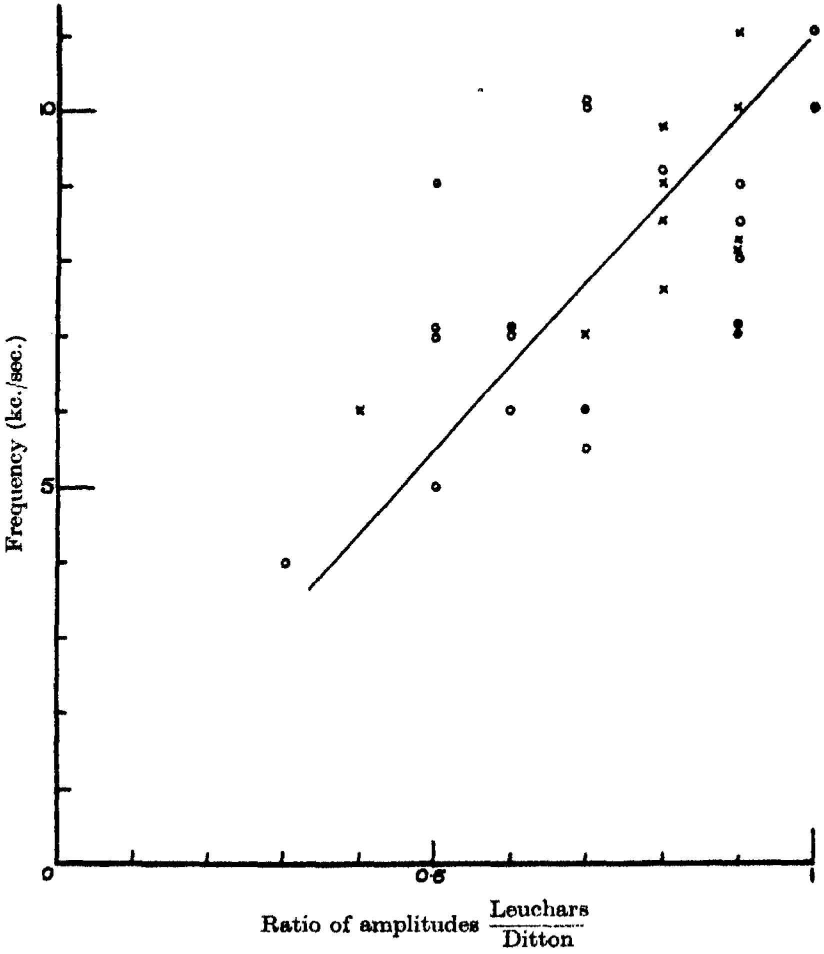


Fig. 8—Scatter diagram showing the manner in which the ratio of peak amplitudes of the received form at the two stations varies with the predominant frequency.

other factor, presumably the magnetic azimuth of transmission. The trajectories from the Atlantic sources considered to Leuchars are only 4 % longer, but 20° nearer the magnetic meridian than are the trajectories from these sources to Slough.

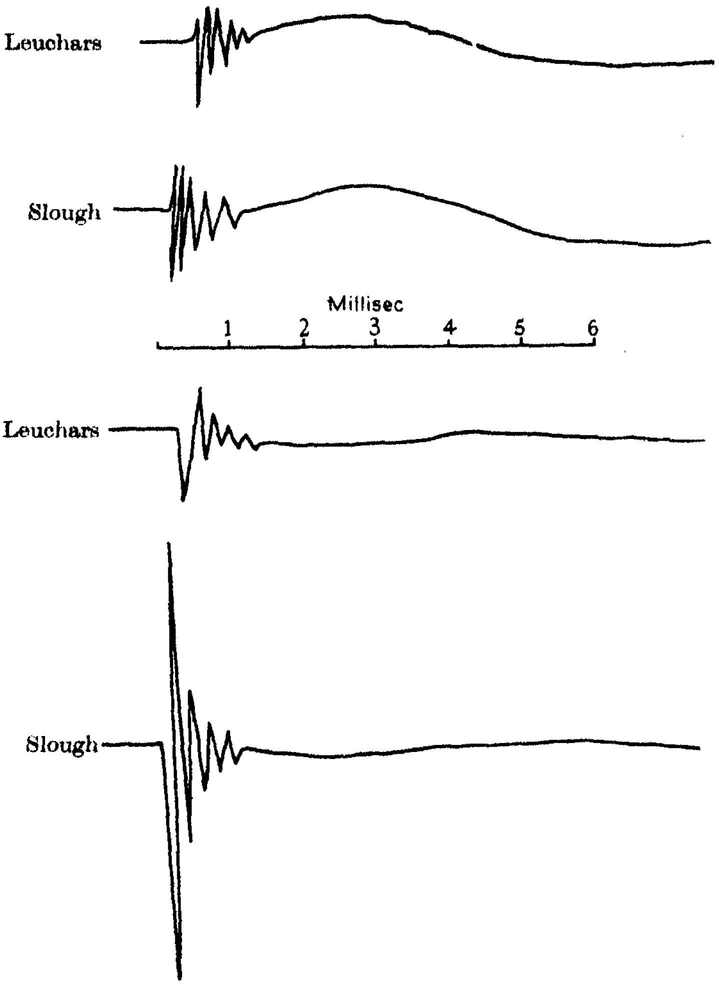


FIG. 9—Tracings from records at Slough and Leuchars of two atmospherics from the same storm centre, showing large variations in the relative amplitudes.

Tracings of the wave forms of two atmospherics from this area as recorded at Slough and Leuchars respectively, and occurring within a few seconds of each other, are shown on common time and amplitude scales in fig. 9. It will be noted that, in these examples, the amplitude of the atmospheric received at Slough, relative to that at Leuchars, was less for the atmospheric with the higher leading frequency, as suggested by fig. 8.

This complicating differential effect has been mentioned as a further indication of the variety of factors intervening to produce, from an established diversity of forms at origin, a still greater diversity of forms at the observing stations. It offers, in fact, a convenient occasion for the general remark that all the quantitative data given in the present paper must be treated as the product of insufficient sampling by processes of insufficient resolving power, and so, as indicative of trends and orders of magnitude only. Any attempt to explore or explain the minor details of the curves shown would, in the authors' view, be uneconomical, since the improvements in technique now being effected offer the prospect of much more satisfactory observational material in the near future. It should, however, be said, on the other hand, that such discussion as is here given in broad outline is believed to be statistically justified; we are, in fact, anxious to make available just so much of the broad outline as is reasonably established, and to avoid any appearance of a precision which the distribution and the technique do not justify.

Reverting to the limited and simplified summary of the Slough observations alone in relation to distance from source as determined by Slough and Leuchars jointly, we present in fig. 10 a series of curves relating the mean values of important parameters to distance from source. It will be seen *from these figures that the time interval between the start of the oscillatory* component and that of the slow component varies from about -1 msec., sometimes reaching -3 msec. near the source (the oscillatory component in this case starting at this time interval after the start of the slow component), through zero at about 500 km. to about $1\frac{1}{2}$ msec. (the high frequencies now leading) at about 3000 km.

The oscillatory component, which usually builds up to its maximum amplitude in one or two oscillations, and decays to negligible amplitude after some six or more oscillations, has predominant quasi-frequencies of 5 or 6 kc./sec. at its head up to 300 km., and the predominant and leading frequency increases to 10 kc./sec. at about 2500 km. The peak amplitude of the oscillatory component decays with distance from values of the order of 500 mV/m. at 100 km. to 200 mV/m. at 1000 km. and to 25 mV/m. at 4000 km.

The slow component, which has a peak amplitude of 125 mV/m. at 100 km., decays to 50 mV/m. at 500 km. and to some 10 mV/m. at 4000 km. The duration of its first "half-cycle" appears to decrease from about 2·5 msec. at 100 km. to a minimum of 1·4 msec. at 300 km. Thence it increases

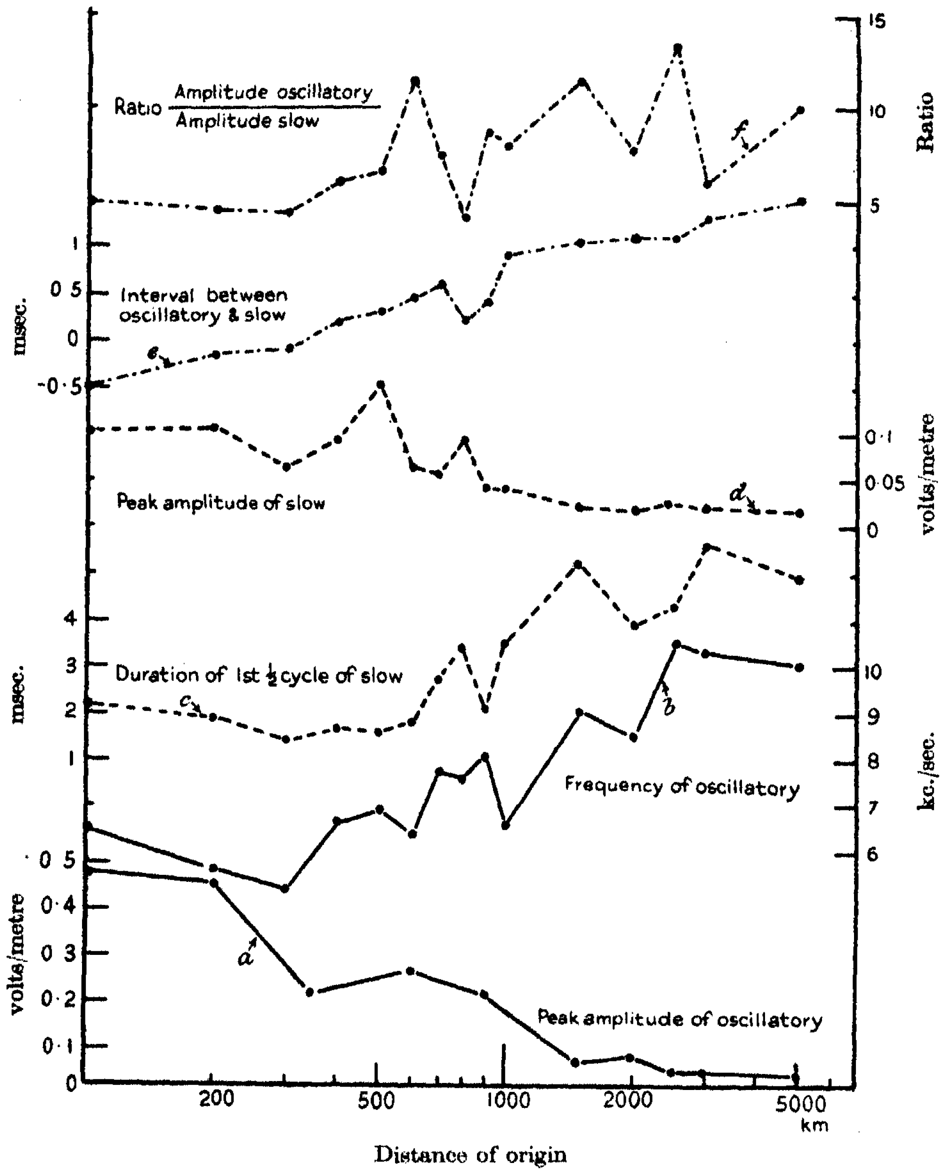


FIG. 10—Curves showing variation with distance of origin of *a*, peak amplitude of oscillatory component; *b*, frequency of oscillatory component; *c*, duration of first half-cycle of slow component; *d*, peak amplitude of slow component; *e*, interval between oscillatory and slow; *f*, ratio-amplitude of oscillatory to amplitude of slow.

to some 3·5 msec. at 1000 km. and to about 4·5 msec. at 2000 km. It is reasonably strongly indicated that at 4000 km. from the source the mean duration is still increasing, from its value of about 6 msec. there, at a rate of the rough order of + 1 msec. per 1000 km.

The ratio of amplitudes of the oscillatory to the slow portion of the whole atmospheric thus falls off with distance, but the oscillatory component is at all normal distances of greater amplitude than is the slow. It is, however, possible to trace in some of the groupings of field changes (each group emanating, so far as can reasonably be inferred, from the discharge processes at a single source) much smaller oscillatory components preceding the main one. One example may be seen in fig. 3*c*, Plate 11. This precursor is usually about one-tenth of the main oscillatory component in amplitude, and precedes it by an interval of the order of a millisecond. The quasi-frequencies contained in it are still higher than those already discussed, and are not, in fact, adequately resolved at a film speed of 1 m./sec. Further attention is being devoted to these higher quasi-frequencies; meanwhile a specimen record at a film speed of 10 m./sec. is shown in fig. 11 to illustrate the nature of this precursor. It is not possible to exclude completely the hypothesis that this further separation is enhanced, or even produced, by the differences in group velocity with frequency which separate the main oscillatory components from the slow form. It is, however, very much more probable that the precursors are radiated during those minor discontinuities in the (*a*) portion of the discharge (Appleton and Chapman 1937, p. 19) which have already been associated with the discontinuities in the leader stroke identified by Schonland, Malan and Collens (1935).

About one in every twenty of the oscillatory forms recorded may be found to resemble No. *b* of fig. 3, Plate 11. It would appear that in such cases the originating discharge takes the form of some twenty or thirty component discharges each radiating a decaying oscillatory train similar to the typical oscillatory component discussed above. This whole sequence appears to be associated with a single slow form and occupies a time interval of a few milliseconds only. In the case of these special forms again a higher photographic resolution is required than that permitted by the simple 1 m./sec. film recording, and further studies at higher film speeds will be discussed in a subsequent paper of this series.

The work described was carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research. We are indebted to our colleagues in the Radio Department of the National Physical Laboratory for much assistance; we wish to mention in

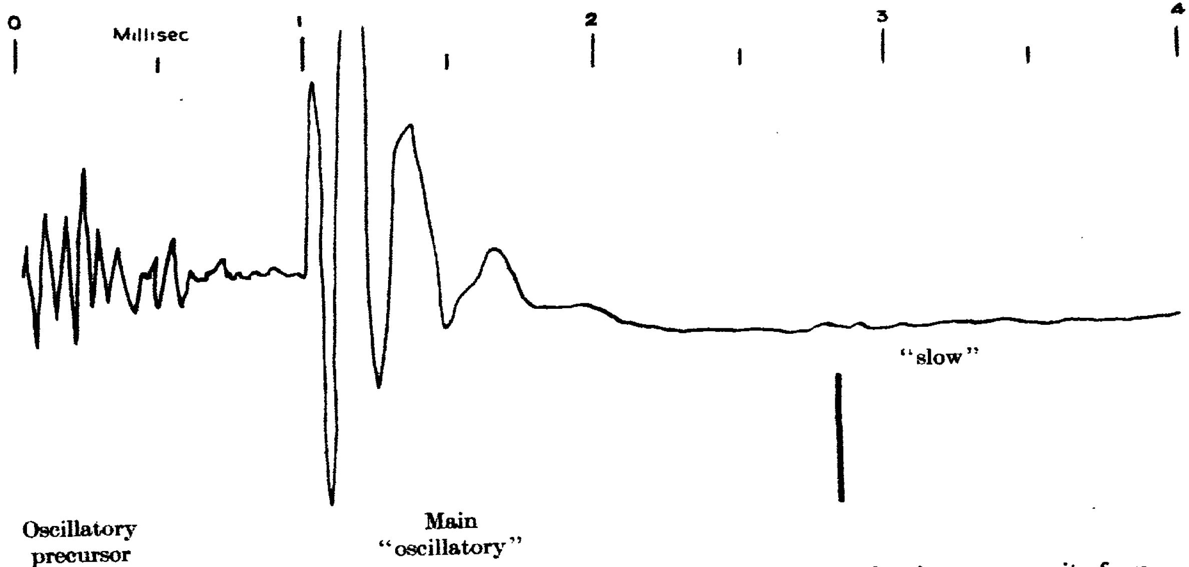


FIG. 11—Tracing from a record made at a speed of 10 metres per sec. showing a composite form containing oscillatory and slow components, preceded by a form of still higher frequency but of much smaller amplitude.

particular Mr. G. H. Calver, who carries out, single-handed, the work of the Leuchars observing station.

In relation to the material summarized in Tables I to VI we are indebted to many authorities and individuals for special facilities and co-operation. The work resulting in Table I was done in association with the Admiralty, and Messrs. C. E. Horton and J. W. Drabble of H.M. Signal School co-operated in the experimental and observational work. The work leading to Table III could not have been carried out without the generous co-operation of the Admiralty, the War Office, and the Physical Department of the Government of Egypt. Particular thanks are due to Dr. H. Knox-Shaw, then Director of Helwan Observatory, and to Major (now Colonel) J. P. G. Worlledge, Royal Corps of Signals, and his officers. The work summarized in Table V was made possible by the cordial co-operation of the Government of the Sudan, and special thanks are due to Bimbashi H. Williams of the Department of Posts and Telegraphs at Khartoum.

SUMMARY

The results of eye-and-hand delineations of atmospheric wave forms observed in the Red Sea and Indian Ocean, Egypt, the Anglo-Egyptian Sudan and in "control" observations in south-east England are presented in tabular form.

Photographic recording of wave form and of apparent direction of arrival at the two ends of a 560 km. base-line reveals the details of the dispersive process by which the higher frequency components, travelling with a higher group velocity and subject to heavier attenuation than the lower frequency components, form, at distances over some 500 km., a discrete oscillatory component preceding and completely detached from a slow form of the type delineated by the eye-and-hand method. The principal characteristics of the resulting dual wave form are plotted against distance over ranges up to 4000 km.

The oscillatory component is shown to contain prominent frequencies of the order of 5–20 kc./sec., and to have peak amplitudes falling from 500 mV/m. at 100 km. to 25 mV/m. at 4000 km. The slow form, with peak amplitudes of 125 mV/m. at 100 km. and 10 mV/m. at 4000 km. has a first "half-cycle" whose duration lies between 1·5 and 2·5 msec. in the first few hundred kilometres and rises to 6 msec. at 4000 km.

Some of the complexities in the factors controlling relative amplitudes and structure at the two ends of the base-line are briefly indicated.

The occurrence of a "precursor" of small amplitude, comprising still

higher frequency components, and approximately a millisecond in advance of the main oscillatory component, is provisionally associated with the discontinuities in the leader stroke identified by Schonland, Malan and Collens.

REFERENCES

Appleton, E. V. and Chapman, S. 1937 *Proc. Roy. Soc.* A, **158**, 1.

Appleton, E. V., Watt, R. A. Watson and Herd, J. F. 1926 *Proc. Roy. Soc.* A, **111**, 615.

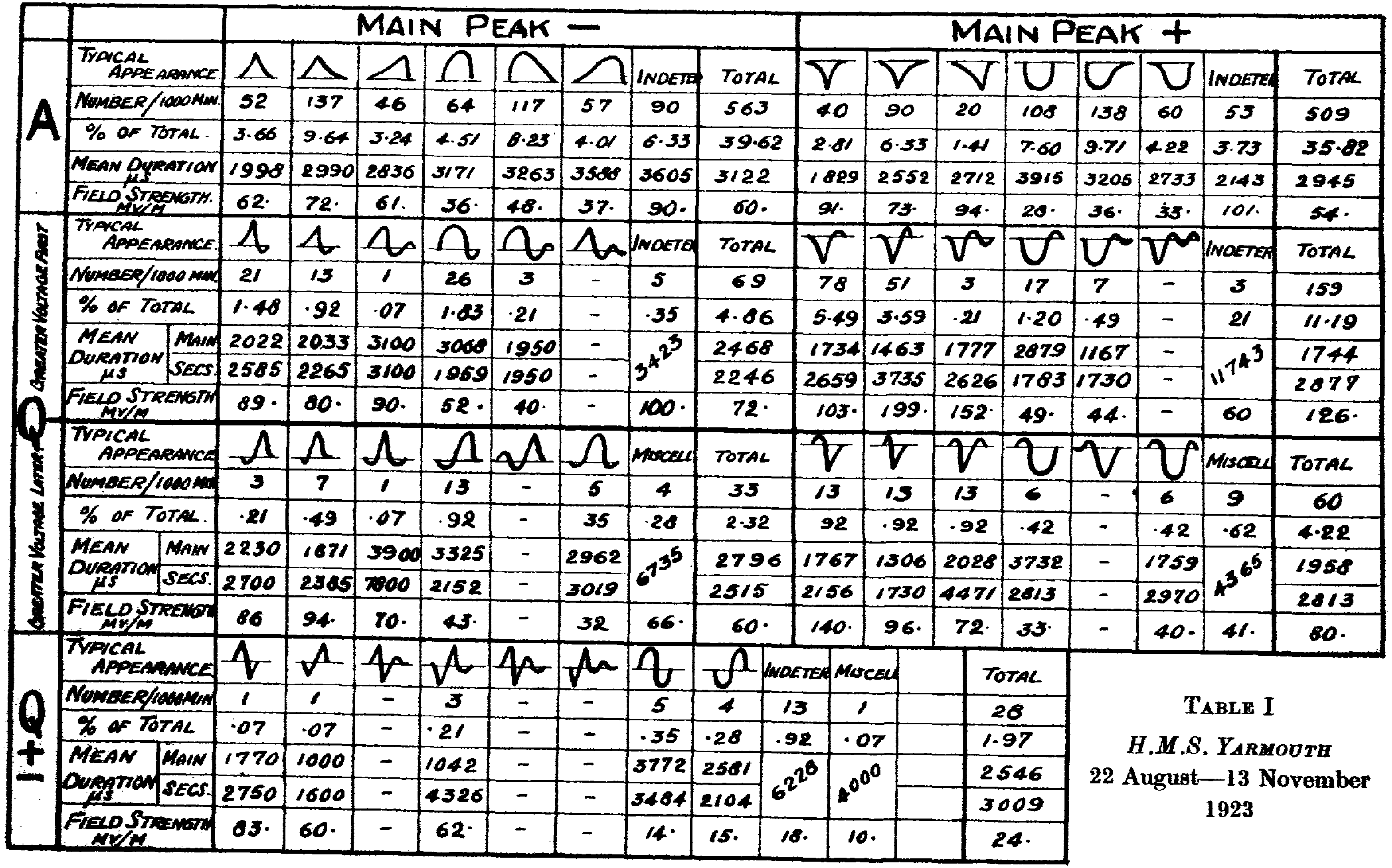
Schonland, B. F. J., Malan, D. J., and Collens, H. 1935 *Proc. Roy. Soc.* A, **152**, 595.

Watt, R. A. Watson and Appleton, E. V. 1923 *Proc. Roy. Soc.* A, **103**, 84.

Watt, R. A. Watson, Herd, J. F. and Bainbridge-Bell, L. H. 1933 "The Cathode Ray Oscillograph in Radio Research." H.M.S.O.

Watt, R. A. Watson 1927 Supplement *Quart. J.R. Met. Soc.* **53**.

— 1929 *Quart. J.R. Met. Soc.* **54** (July).



		MAIN PEAK —								MAIN PEAK +							
A	TYPICAL APPEARANCE							INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN.	52	137	46	64	117	57	90	563	40	90	20	108	138	60	53	509
	% OF TOTAL.	3·66	9·64	3·24	4·51	8·23	4·01	6·33	39·62	2·81	6·33	1·41	7·60	9·71	4·22	3·73	35·82
	MEAN DURATION μS	1998	2990	2836	3171	3263	3588	3605	3122	1829	2552	2712	3915	3206	2733	2143	2945
	FIELD STRENGTH. MV/M	62·	72·	61.	36·	48·	37·	90·	60·	91·	73·	94·	28·	36·	35·	101·	54·
Q GREATER VOLTAGE FIRST	TYPICAL APPEARANCE.							INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN.	21	13	1	26	3	-	5	69	78	51	3	17	7	-	3	159
	% OF TOTAL	1·48	·92	·07	1·83	·21	-	·35	4·86	5·49	3·59	·21	1·20	·49	-	21	11·19
	MEAN DURATION μS MAIN	2022	2033	3100	3068	1950	-	3423	2468	1734	1463	1777	2879	1167	-	11743	1744
	MEAN DURATION μS SECS.	2585	2265	3100	1969	1950	-		2246	2659	3735	2626	1783	1730	-		2877
	FIELD STRENGTH MV/M	89·	80·	90·	52·	40·	-	100·	72·	103·	199·	152·	49·	44·	-	60	126·
GREATER VOLTAGE LATER	TYPICAL APPEARANCE							MISCELL	TOTAL							MISCELL	TOTAL
	NUMBER/1000 MIN.	3	7	1	13	-	5	4	33	13	13	13	6	-	6	9	60
	% OF TOTAL.	·21	·49	·07	·92	-	35	·28	2·32	92	·92	·92	·42	-	·42	·62	4·22
	MEAN DURATION μS MAIN	2230	1671	3900	3325	-	2962	6735	2796	1767	1306	2028	3732	-	1759	4365	1958
	MEAN DURATION μS SECS.	2700	2385	7800	2152	-	3019		2515	2156	1730	4471	2813	-	2970		2813
	FIELD STRENGTH MV/M	86	94·	70·	43·	-	32	66·	60·	140·	96·	72·	33·	-	40·	41·	80·

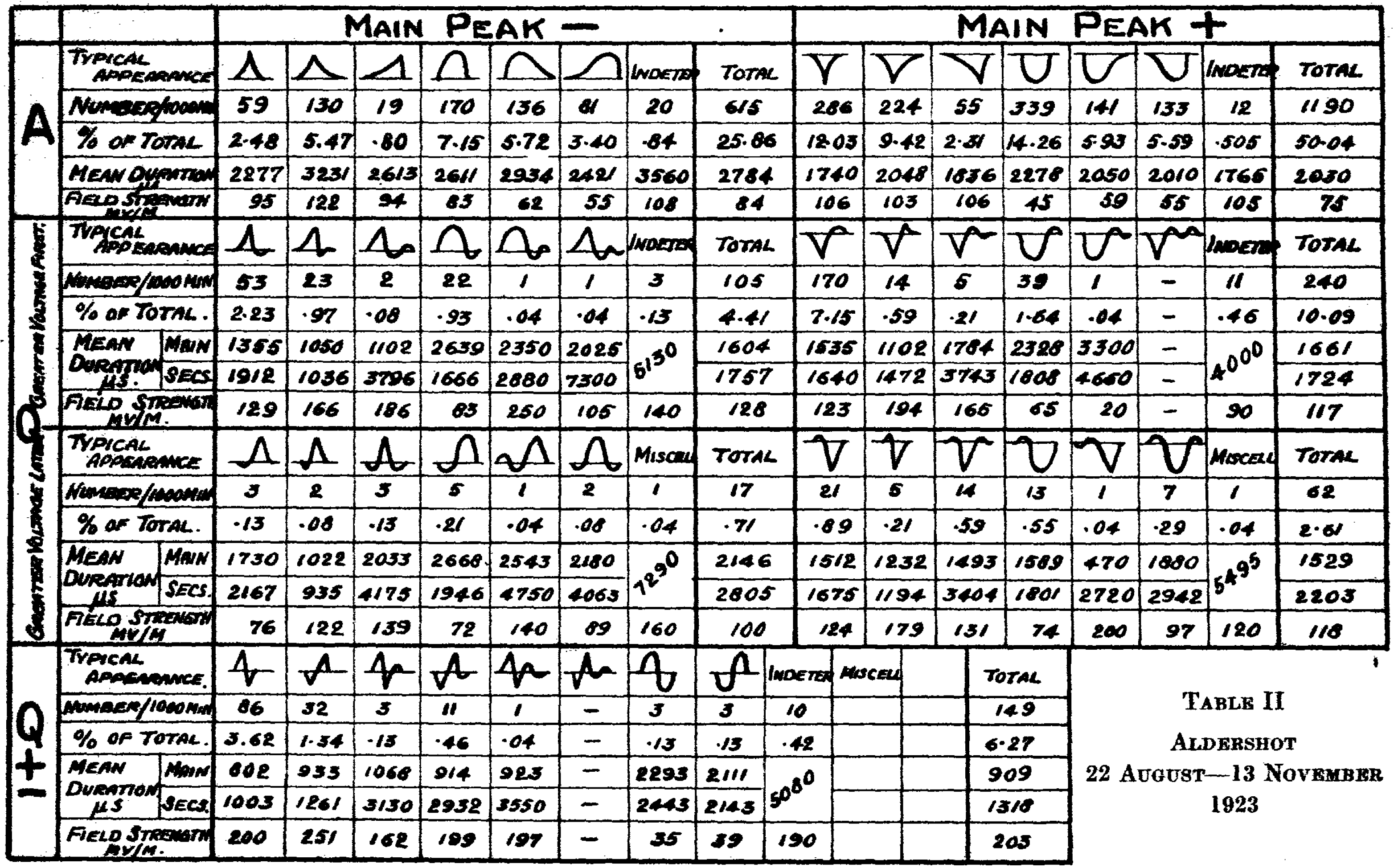
I+Q	TYPICAL APPEARANCE									INDETER	MISCELL		TOTAL
	NUMBER/1000 MIN	1	1	-	3	-	-	5	4	13	1		28
	% OF TOTAL	·07	·07	-	·21	-	-	·35	·28	·92	·07		1·97
	MEAN DURATION μS MAIN	1770	1000	-	1042	-	-	3772	2581	6228	4000		2546
	MEAN DURATION μS SECS.	2750	1600	-	4326	-	-	3484	2104				3009
	FIELD STRENGTH MV/M	83·	60·	-	62·	-	-	14·	15·	18·	10·		24·

TABLE I

H.M.S. YARMOUTH

22 August—13 November

1923

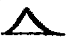
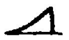
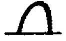
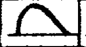
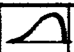
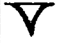
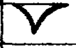
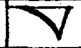
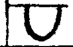
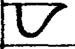
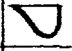
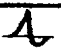
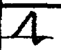
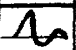
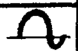
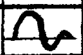
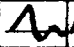
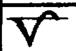
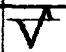
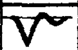
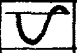
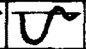
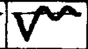
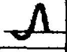
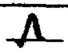
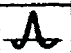
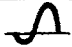
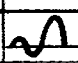
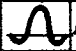
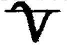
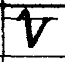
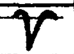
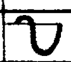
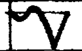
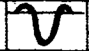


		MAIN PEAK —								MAIN PEAK +							
A	TYPICAL APPEARANCE							INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN	59	130	19	170	136	81	20	615	286	224	55	339	141	133	12	1190
	% OF TOTAL	2·48	5.47	·80	7·15	5·72	3·40	·84	25·86	12·03	9·42	2·31	14·26	5·93	5·59	·505	50·04
	MEAN DURATION µS	2277	3231	2613	2611	2934	2421	3560	2784	1740	2048	1836	2278	2050	2010	1766	2030
	FIELD STRENGTH MV/M	95	122	94	83	62	55	108	84	106	103	106	45	59	55	105	75
Q GREATER VOLTAGE FIRST.	TYPICAL APPEARANCE							INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN	53	23	2	22	1	1	3	105	170	14	5	39	1	–	11	240
	% OF TOTAL.	2·23	·97	·08	·93	·04	·04	·13	4·41	7·15	·59	·21	1·64	·04	–	·46	10·09
	MEAN DURATION µS. MAIN	1355	1050	1102	2639	2350	2025	6130	1604	1535	1102	1784	2328	3300	–	4000	1661
	MEAN DURATION µS. SECS.	1912	1036	3796	1666	2880	7300		1757	1640	1472	3743	1808	4660	–		1724
	FIELD STRENGTH MV/M.	129	166	186	83	250	105	140	128	123	194	166	65	20	–	90	117
Q GREATER VOLTAGE LATER	TYPICAL APPEARANCE							MISCELL	TOTAL							MISCELL	TOTAL
	NUMBER/1000 MIN	3	2	3	5	1	2	1	17	21	5	14	13	1	7	1	62
	% OF TOTAL.	·13	·08	·13	·21	·04	·08	·04	·71	·89	·21	·59	·55	·04	·29	·04	2·61
	MEAN DURATION µS MAIN	1730	1022	2033	2668	2543	2180	7290	2146	1512	1232	1493	1589	470	1880	5495	1529
	MEAN DURATION µS SECS.	2167	935	4175	1946	4750	4063		2805	1675	1194	3404	1801	2720	2942		2203
	FIELD STRENGTH MV/M	76	122	139	72	140	89	160	100	124	179	131	74	200	97	120	118

1+Q	TYPICAL APPEARANCE.									INDETER	MISCELL		TOTAL
	NUMBER/1000 MIN	86	32	3	11	1	—	3	3	10			149
	% OF TOTAL.	3.62	1·34	·13	·46	·04	—	·13	·13	·42			6·27
	MEAN DURATION µS MAIN	802	933	1068	914	923	—	2293	2111	5080			909
	MEAN DURATION µS SECS.	1003	1261	3130	2932	3550	—	2443	2143				1318
	FIELD STRENGTH MV/M.	200	251	162	199	197	—	35	39	190			203

TABLE II

ALDERSHOT
22 AUGUST—13 NOVEMBER
1923

		MAIN PEAK −								MAIN PEAK +							
A	TYPICAL APPEARANCE.							INDETER.	TOTAL							INDETER.	TOTAL
	NUMBER/1000 MIN.	75	227	66	186	475	168	121	1318	180	418	197	373	594	644	113	2519
	% OF TOTAL.	1·63	4·93	1·43	4·04	10·31	3·65	2·62	28·61	3·91	9·08	4·28	8·10	12·90	13·98	2·453	54·69
	MEAN DURATION μS	1830	2459	1862	2130	2670	2170	–	2365	1330	1512	1534	1985	1710	1645	–	1627
	FIELD STRENGTH MV/M.	67·0	89·0	94·3	78·2	78·2	70·0	160·0	86·6	94·8	112·2	89·9	51·8	74·1	68·8	153·5	81·8
Q GREATER VOLTAGE FIRST.	TYPICAL APPEARANCE							INDETER.	TOTAL							INDETER.	TOTAL
	NUMBER/1000 MIN.	30	22	1	33	2	1	4	93	171·	112	7	70	2	1	8	371
	% OF TOTAL.	·65	·48	·02	·72	·04	·02	·09	2·02	3·71	2·43	·15	1·52	·04	·02	·18	8·05
	MEAN DURATION μS MAIN	1676	1661	500	1863	2100	2000	2130	1754	1410	1170	1475	2095	1605	1070		1469
	SECS.	1664	1257	2125	1133	2876	3830	2780	1480	1690	1885	2650	1300	1695	2125		1695
	FIELD STRENGTH MV/M.	122·6	100·3	133·0	77·5	69·0	145·0	132·0	101·1	125·5	185·0	176·0	71·1	135·5	333·0	315·0	138·9
GREATER VOLTAGE LATER.	TYPICAL APPEARANCE							MISCELL.	TOTAL							MISCELL.	TOTAL
	NUMBER/1000 MIN.	11	12	5	23	1	1	1	54	31	16	7	36	2	3	11	106
	% OF TOTAL	·24	·26	·11	·50	·02	·02	·02	1·17	·67	·55	·15	·78	·04	·07	·24	2·30
	MEAN DURATION μS MAIN	1517	2283	1895	1727	650	2425	875	1822	1485	1255	869	1685	950	2055	2848	1484
	SECS.	1923	1337	3376	1407	1350	1700		1710	1300	949	2602	1180	2300	2532		1360
	FIELD STRENGTH MV/M	68·5	93·2	75·0	64·3	240·0	47·5	350·0	77·7	86·0	104·1	232·0	47·2	189·0	70·0	321·0	111·1

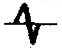
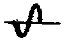
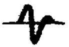
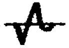
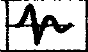
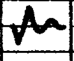
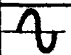
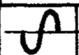
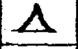
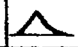
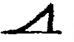
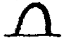
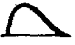
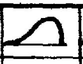
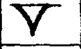
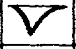
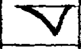
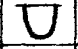
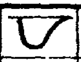
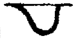
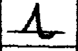
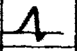
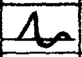
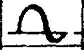
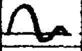
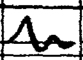
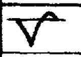
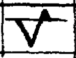
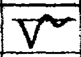
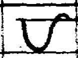
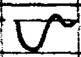
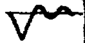
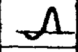
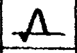
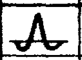
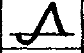
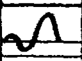
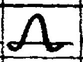
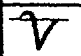
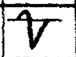
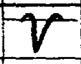
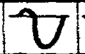
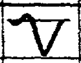
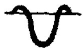
Q±	TYPICAL APPEARANCE.									INDETER.	MISCELL.		TOTAL
	NUMBER/1000 MIN.	19	14	6	2	1	–	32	53	18	–		145
	% OF TOTAL.	·41	·31	·13	·04	·02	–	·70	1·15	·39	–		3·15
	MEAN DURATION μS MAIN	1210	1365	560	320	300	–	1745	2065	–	–		1671
	SECS.	1090	1315	3000	2745	1380	–	1630	1745	–	–		1641
	FIELD STRENGTH MV/M	73·0	57·2	133·0	127·0	118·0	–	30·9	34·0	319·3	–		82·4

TABLE III

HELWAN

15 DECEMBER 1923—
15 MARCH 1924

			MAIN PEAK —								MAIN PEAK +							
A	TYPICAL APPEARANCE								INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN		31	148	27	190	293	114	2	805	42	106	33	328	350	153	–	1011
	% OF TOTAL		1·43	6·83	1·25	8·76	13·51	5·26	·09	37·13	1·94	4·89	1·50	15·12	16·14	7·05	–	46·63
	MEAN DURATION µS		1436	2855	2028	2765	2719	2109	–	2596	1756	2058	1741	2332	2338	1901	–	2197
	FIELD STRENGTH MV/M		122·0	134·2	97·5	32·5	53·3	47·7	160·0	66·9	51·6	64·7	57·2	23·6	32·9	32·4	–	34·7
Q GREATER VOLTAGE FIRST	TYPICAL APPEARANCE								INDETER	TOTAL.							INDETER	TOTAL.
	NUMBER/1000 MIN		49	12	2	43	1	1	1	109	45	5	1	56	–	–	1	108
	% OF TOTAL		2·26	·55	·09	2·00	·02	·05	·05	5·03	2·07	·23	·05	2·61	–	–	·02	4·98
	MEAN DURATION µS	MAIN.	1582	786	1854	2181	780	2350	–	1742	1584	1124	1560	2155	–	–	–	1862
		SECS.	1802	1039	4008	1619	1750	5280	–	1726	1823	2123	3120	1957	–	–	–	1916
	FIELD STRENGTH MV/M		142·2	194·2	182·5	71·7	160·0	135·0	37·5	119·9	85·5	160·5	105·0	43·7	–	–	55·0	67·2
GREATER VOLTAGE LATER	TYPICAL APPEARANCE.								MISCELL	TOTAL							MISCELL	TOTAL
	NUMBER/1000 MIN		5	1	1	10	1	6	1	25	9	4	2	10	1	3	1	30
	% OF TOTAL		·23	·05	·05	·46	·03	·28	·05	1·15	·41	·18	·09	·46	·05	·14	·05	1·38
	MEAN DURATION µS	MAIN.	1879	1870	4539	2564	580	1493	–	2191	1370	996	1819	1517	780	1826	–	1421
		SECS.	2163	1250	4177	1424	970	1739	–	1779	1494	570	3727	1466	2920	2087	–	1612
	FIELD STRENGTH MV/M		96·2	160·0	149·3	56·6	120·0	137·9	200·0	98·4	64·9	135·0	69·5	47·2	66·0	74·6	135·0	71·5

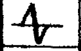
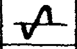
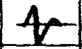
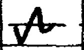
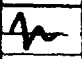
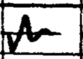
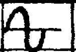
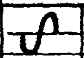
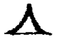
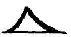
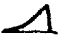
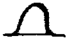
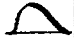
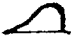
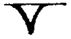
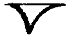
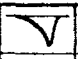
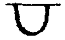
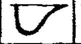
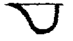
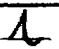
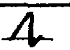
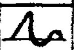
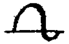
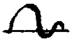
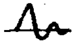
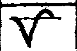
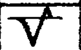
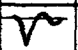
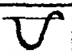
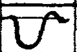
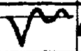
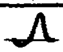
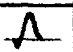
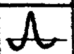
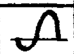
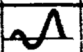
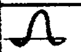
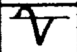
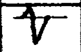
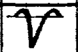
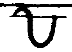
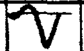
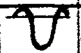
Q+−	TYPICAL APPEARANCE										INDETER	MISCELL		TOTAL
	NUMBER/1000 MIN		27	27	1	7	–	1	8	8	2	–		80
	% OF TOTAL		1·24	1·24	·05	·32	–	·02	·37	·37	·09	–		3·70
	MEAN DURATION µS	MAIN.	672	600	666	770	–	200	3469	1806	–	–		1051
		SECS.	899	741	2807	2500	–	780	2459	1946	–	–		1284
	FIELD STRENGTH MV/M.		161·1	217·9	235·0	219·0	–	165·0	23·6	23·3	25·2	–		153·5

TABLE IV

ALDERSHOT
15 DECEMBER 1923—
15 MARCH 1924

		MAIN PEAK —								MAIN PEAK +							
A	TYPICAL APPEARANCE							INDETER.	TOTAL							INDETER.	TOTAL
	NUMBER/1000 MIN	113	269·5	138·5	157·5	389	123·5	98·5	1289·5	152·5	399	123·5	245	456	226·5	93 3 96·8	1699
	% OF TOTAL	2·82	6·71	3·45	3·92	9·69	3·08	2·45	32·13	3·80	9·94	3·08	6·10	11·36	5·64	2·32 ·09 2·41	42·33
	MEAN DURATION µS	1196	1676	1376	1932	2210	1432	–	1756	694	858	921	1973	1467	1343	–	1257
	FIELD STRENGTH MV/M	100·6	116·3	82·7	54·0	63·7	75·3	87·6	81·8	193·9	151·3	121·0	54·7	66·3	67·9	[illegible]	105·2
Q GREATER VOLTAGE FIRST	TYPICAL APPEARANCE							INDETER.	TOTAL							INDETER.	TOTAL
	NUMBER/1000 MIN	51	22	1·5	41	–	1·5	–	117	249·5	286·5	8	39	–	1·5	10·5	595
	% OF TOTAL	1·27	·55	·04	1·02	–	·04	–	2·92	6·21	7·14	·20	·97	–	·04	·26	14·82
	MEAN DURATION µS. MAIN	1256	945	800	1982	–	2000	–	1456	727	640	703	1419	–	100	–	729
	SECS.	1737	880	2000	1645	–	8000	–	1627	1647	1791	2805	776	–	600	–	1672
	FIELD STRENGTH MV/M	127·2	155·2	181·0	110·4	–	36·0	–	126·1	235·2	299·6	210·5	98·2	–	141·0	543·0	262·1
GREATER VOLTAGE LATER	TYPICAL APPEARANCE							MISCELL.	TOTAL							MISCELL.	TOTAL
	NUMBER/1000 MIN	21	12	4·5	7	1·5	3·5	1·5	51	5	5	15·5	8	–	3	116·5	153
	% OF TOTAL	·52	·30	·11	·17	·04	·09	·04	1·27	13	13	·38	·20	–	·07	2·90	3·81
	MEAN DURATION µS MAIN	802	1431	1467	2114	4000	400	–	1269	633	666	401	3220	–	2900	–	1292
	SECS	1310	1223	2234	2543	4000	450	–	1568	733	500	1192	1600	–	2850	–	1260
	FIELD STRENGTH MV/M.	89·1	99·4	137·0	40·8	36·0	167·0	135·0	94·2	79·7	136·0	246·9	74·0	–	126·0	310·0	274·5

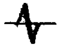
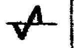
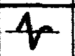
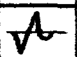
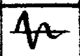
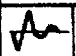
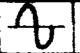
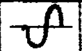
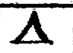
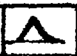
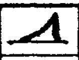
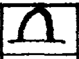
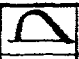
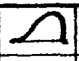
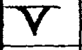
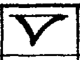
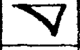
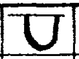
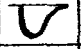
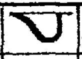
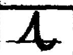
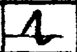
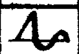
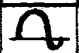
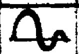
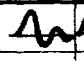
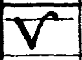
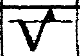
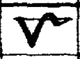
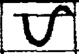
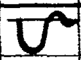
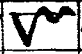
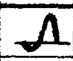
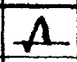
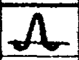
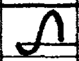
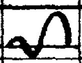
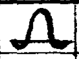
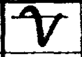
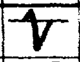
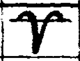
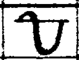
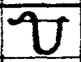
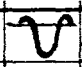
Q+−	TYPICAL APPEARANCE									INDETER.	MISCELL.		TOTAL
	NUMBER/1000 MIN	5	8	3	15	–	–	31·5	32·5	8	20		109·5
	% OF TOTAL	·13	·20	·07	·04	–	–	·78	·81	·20	·50		2·73
	MEAN DURATION µS MAIN	1175	561	175	500	–	–	1565	2012	–	–		1550
	SECS.	1875	805	1400	2000	–	–	1903	1781	–	–		1665
	FIELD STRENGTH MV/M	239·5	53·2	86·0	127·0	–	–	29·5	24·5	159·7	101·5		64·9

TABLE V

KHARTOUM

1 APRIL—30 APRIL 1924

			MAIN PEAK —								MAIN PEAK +							
A	TYPICAL APPEARANCE								INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN		37	77	7	417	318·5	125	4·5	986	53·5	91	20	631	427·5	239	2·5	1464·5
	% OF TOTAL.		1·30	2·72	·25	14·71	11·23	4·41	·16	34·78	1·89	3·21	·71	22·26	15·08	8·43	·08	51·66
	MEAN DURATION µS		1127	2301	1634	2790	3015	2337	1050	2688	1269	1643	1924	2371	2843	2289	985	2401
	FIELD STRENGTH MV/M		197·9	189·4	118·4	32·4	54·1	40·0	61·6	59·6	127·0	151·7	111·6	27·2	32·8	34·0	165·0	42·8
Q GREATER VOLTAGE FIRST.	TYPICAL APPEARANCE								INDETER	TOTAL							INDETER	TOTAL
	NUMBER/1000 MIN		8·5	8·5	1·5	39	–	1·5	4	63	23·5	14·5	–	49·5	1·5	–	–	89
	% OF TOTAL		·30	·30	·05	1·38	–	·05	·14	2·22	·83	·51	–	1·75	·05	–	–	3·14
	MEAN DURATION µS	MAIN	1833	513	150	3593	–	2500	4388	2882	1472	510	–	2966	1950	–	–	2154
		SECS.	1618	837	350	1742	–	6250	4388	1845	2756	1470	–	2618	4290	–	–	2496
	FIELD STRENGTH MV/M		121·6	160·7	400·0	50·2	–	80·0	99·7	86·9	202·6	288·4	–	73·2	35·0	–	–	141·8
Q GREATER VOLTAGE LATER.	TYPICAL APPEARANCE								MISCELL	TOTAL							MISCELL	TOTAL
	NUMBER/1000 MIN		7	1·5	1·5	17	–	11·5	–	38·5	1·5	1·5	1·5	21	–	4·5	–	30
	% OF TOTAL		·25	·05	·05	·60	–	·41	–	1·36	·05	·05	0·5	·74	–	·17	–	1·06
	MEAN DURATION µS	MAIN	1923	300	390	2795	–	3585	–	2682	300	580	620	2766	–	2420	–	2374
		SECS.	2440	300	1550	1950	–	3749	–	2488	500	580	3740	2255	–	3690	–	2373
	FIELD STRENGTH MV/M		113·0	585·0	95·0	53·0	–	62·6	–	89·1	265·0	535·0	35·0	53·8	–	96·7	–	93·9

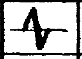
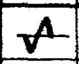
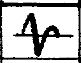
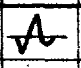
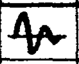
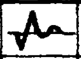
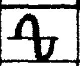
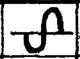
Q+−	TYPICAL APPEARANCE										INDETER	MISCELL	MISCELL	TOTAL
	NUMBER/1000 MIN		30·5	86·5	2·5	21	–	–	2·5	1·5	2·5	1·5 17	15·5	164
	% OF TOTAL		1·09	3·06	·08	·74	–	–	·08	·05	·08	·05 ·60	·55	5·78
	MEAN DURATION µS	MAIN	729	483	300	559	–	–	4135	3120	–	1560	–	634
		SECS.	1169	745	2810	2868	–	–	2420	3120	–		–	1232
	FIELD STRENGTH MV/M		177·7	208·0	300·0	185·7	–	–	17·5	20·0	310·0	100·0 213·1	224·1	198·4

TABLE VI

ALDERSHOT

1 APRIL—30 APRIL 1924

The Behaviour of an Osglim Lamp
I—Introduction, Drift, Equilibrium and A.C. Bridge Measurements

By Lewis F. Richardson, F.R.S.

(*Received* 12 *January* 1937)

[Plate 13]

1—Introduction

Fig. 1 shows the relation between the voltage across the osglim and the current through it during various periodic disturbances. We see a tangle of curves intersecting each other, and one of them even intersecting itself. Evidently there are many varieties of behaviour to be discussed. An attempt will be made to formulate a single law that shall apply to them all. Briefly, the first clue is that each loop except the self-intersection may be regarded as the deformation of a circle and that the equilibrium states, shown by the dotted line in fig. 1, are the deformation of a diameter of that circle.

The experiments here described, although in themselves plain physics, have been pursued in the hope that the resulting empirical equations may afterwards be applied, with due caution, to some of the more mechanistic parts of psychology. How this may be is explained by quoting the summary of a paper in the *Psychological Review* (Richardson 1930):

"This is a development from an analogy between sensations and spark published by McDougall in *Brain* of 1901. The strain of expectation, which we feel when we wait for the solution of a problem, is here compared with the electric intensity in a spark-gap before the spark passes. The time of imageless suspense is compared with the spark-delay, noting that each of them is fickle. The least intensity with which we must attempt any problem if we are ever to solve it, is compared with the minimum sparking potential. Cleverness, in J. C. M. Garnett's sense (1919*a*, *b*), is compared with the ionization of the spark-gap by X-rays or γ-rays. Having found nine analogies, of which the foregoing are a sample, the author then predicted, from McDougall's theory of inhibition by drainage, that it would be possible to connect two spark-gaps so that the passage of one spark should inhibit the other; the change-over either (*a*) requiring an extra stimulus, as in purposive thinking; or (*b*) occurring automatically, as in reverie. Both predictions have been verified by using neon lamps as spark-gaps.* Altogether sixteen mutually consistent analogies have been collected. These analogies show that when physicists shall have agreed among themselves about the theory of sparking, their formulae will be of great interest to both physiologists and psychologists. For the physical equations will provide a description of the analogous psychological events on giving the following new names to the symbols: Intensity of trying-to-think for potential difference, in-

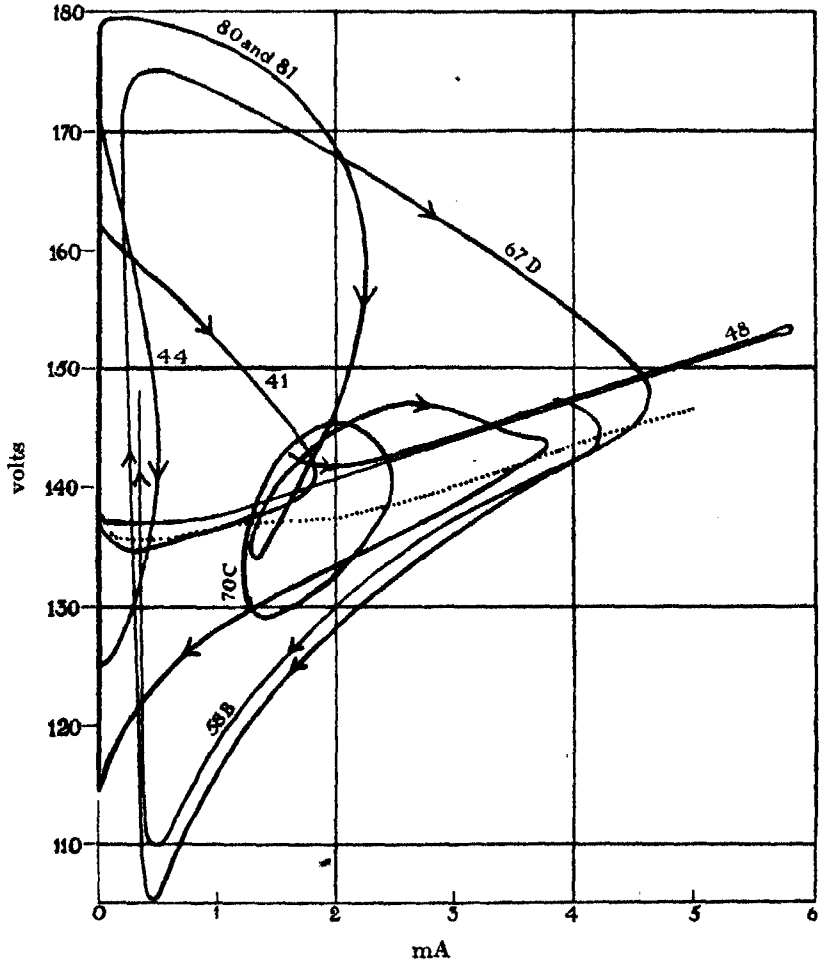


Fig. 1—Various periodic disturbances.†

---

* The experiment (*a*) was exhibited at the Physical and Optical Societies' Exhibition, London, January 1932.

† [10 August 1937. The two uppermost arcs require further correction for strays.]

tensity of imagery for either electric current or energy radiated, cleverness for ionization due to X- or γ-rays."

Mention should also be made of a remarkable electrical model of the heart constructed by van der Pol and van der Mark (1928) from neon lamps and explained by them in connexion with the former's theory of relaxation oscillations.

Psychological analogies seem to be peculiarly liable to misinterpretation. The analogy between mental images and sparks is of the same general type as that discovered by Prandtl when he showed that the equations, which describe the form of a soap film, resemble those which describe the distribution of stress in a twisted metal bar. Griffith and Taylor (1925) have used that analogy to find the stresses in bars of various sections by measuring soap films. How absurd it would be to interpret Prandtl's analogy as meaning that a metal bar consisted of soap solution! Correspondingly the analogy between mental images and sparks must be understood in a mathematical, not in a materialistic sense.

[*Note added* 20 *April* 1937. The author feels bound to sympathize with any reader who should complain that the psychological analogy lacks the clear-cut precision characteristic of physics. A referee suggested that the statement of the analogy should either be omitted or else further developed. Yet to omit any reference to the analogy would be to conceal the true purpose of the investigation, leaving it open to the objection that it has not been planned as industrial nor as atomic physics. On the other hand, to develop the analogy at the present stage would be to ignore some very relevant experiments which are being prepared for publication in a later paper.]

This being the purpose, it seemed better to use the commercial osglim lamp, which was known to give these interesting effects, rather than to try pure neon and electrodes of simple geometric form, as anyone interested mainly in atomic structure would have preferred to do.

[*Note added* 20 *April* 1937. In the earlier stages of this research eleven osglim lamps were used miscellaneously. But the small differences in their behaviour as regards current, voltage and time seemed much less interesting than their agreements. Also the spectra of all the lamps were very similar, showing strong neon and helium lines. Some of them also showed the Balmer lines of hydrogen. But the only connexion found between spectrum and electrical behaviour was that a lamp which exhibited the peculiar flickering glow described by Leyshon (1932) had in its spectrum some faint lines doubtfully attributed to iodine (Richardson 1932). In view of the slight disagreements between different osglims, it seemed best to make all the later experiments on a single specimen. This indeed has been a sufficient task; for it has involved the photographing of hundreds of oscillograms, in addition to the bridge-measurements now reported. The selected lamp was one of those used in 1928 in the original distraction-circuit, and is labelled P. The coil of wire inserted by the makers in the stem was removed at the outset. Unless the contrary is stated, all measurements refer to osglim P with the polarity such that the glow was nearer to the disk electrode than to the "beehive" spiral wire electrode. The spectrum of osglim P has been photographed for reference. It does show hydrogen lines.

Attention has been concentrated on properties intrinsic to the osglim lamp, because once those are discovered its behaviour in an external circuit containing any arrangement of inductances, resistances and capacitances will become predictable.]

Of the many publications concerning the osglim lamp, a few that particularly lead to the present work will be mentioned. The author has had to choose between investigating or reading, so that he owes a general apology to those whose publications he has not read.

The flashing of an osglim lamp was explained by Pearson and Anson (1922 *a*, *b*) by a very simple theory. They assumed that the lamp lit as soon as the voltage across it rose to an "upper critical potential" of about 170 V, and that the lamp ceased to pass current as soon as the voltage across it fell to a "lower critical potential" of about 140 V. They showed that this theory accounted quantitatively for the frequency of flashes of the order of one per sec. It is, however, quite inadequate for general purposes.

Shaxby and Evans (1924) investigated the relation between current and voltage for steady currents less than 0·1 mA, and showed that the upper critical potential of 153 V at zero amperes was probably connected by a continuous curve of equilibrium potential with the lower critical potential of 139 V at 0·1 mA. The connecting curve was not all observed because in one region, where steady currents were expected, flashes occurred. But it was suggested that the flashes might be due to the capacitance of the lamp itself.

Leyshon (1930) by means of a cathode-ray oscillograph, made it clear that interesting effects were to be sought in times briefer than 10^{-4} sec. Leyshon's volt-ampere characteristics showed that flashing lamps passed current at 10 or 15 V below the usual static characteristic. This made one doubt whether a lower critical potential existed.

As time rates of change are in question, a theory by Braunbek is particularly relevant. He follows Townsend, and assumes the ionization by the positive ions to occur in the gas, not at the kathode.

In his "continuity theory" Braunbek (1926) considers a pair of infinitely wide parallel plane electrodes distant s from one another and having an electric intensity E between. If μ_+ and μ_- are the mobilities of the ions and n_+ and n_- the numbers of them per volume, and t is the time,

$$\frac{\partial n_+}{\partial t} = -E\mu_+\frac{\partial n_+}{\partial x} + E\mu_- n_-\alpha + E\mu_+ n_+\beta + N, \tag{1}$$

$$\frac{\partial n_-}{\partial t} = +E\mu_-\frac{\partial n_-}{\partial x} + E\mu_- n_-\alpha + E\mu_+ n_+\beta + N, \tag{2}$$

in which N is the number of ions per time formed in unit volume by external agency, and α and β are the average numbers of ions produced in 1 cm. run by the negative and positive ions respectively.

When the potential is near to its sparking value ϕ_e, the equations lead approximately to

$$\frac{du}{dt} = \frac{\phi}{s}(\mu_+ + \mu_-)\{\epsilon N + 0{\cdot}022uP(\phi - \phi_e)\}, \tag{3}$$

in which u is the electric current,* ϵ is the electronic charge and P is a positive increasing function of ϕ; say $0{\cdot}022\epsilon P = f(\phi)$. More briefly,

$$\frac{du}{dt} = (\text{positive constant})\,\phi\{N + uf(\phi)\,.\,(\phi - \phi_e)\}. \tag{4}$$

In the present experiments N was due only to stray radiation and may be expected to be negligible except at zero current. In these circumstances (4) becomes

$$\frac{1}{u}\frac{du}{dt} = F_1(\phi)\,.\,(\phi - \phi_e), \tag{5}$$

where $F_1(\phi)$ is some unknown function of ϕ.

2—Slow Drift and Recovery

It is comparatively easy to record, with an accuracy of 1 %, the behaviour of the osglim during an interval of 20 sec. But when records made under similar circumstances after an hour, a week or a year have been compared, discrepancies as large as 50 % have occasionally been found. These are certainly not due to experimental error. If they were simply due to the rise of temperature produced by the discharge, the same "relaxation time" would be common to both. But it was observed that a thin thermo-junction, laid against the outside of an osglim (labelled H) and covered with a piece of rubber plaster, attained one-half of its ultimate temperature change in about 2·7 min. when the lamp, complete with makers' ballast, was maintained on 200 V A.C. Some of the drifts in the electrical constants of the osglim have indeed a half-period time of that order; but for others it is very much longer, so that considerable effects persist for a day. The drift is much more rapid and goes further at 5 mA than at 0·1 mA.

---

* The author begs leave to use the symbol u instead of the conventional i for electric current. The advantages of u are threefold. It does not like i confuse with $\sqrt{-1}$. Again $\dot{u}$ can be written for du/dt. Lastly u has no other very common meaning in electrical theory.

Unfortunately, a great many observations had been taken before the importance of drift was fully realized. And even when it is realized there remains the difficulty of defining a standard condition which shall be convenient for all purposes. The problem will be taken up again in the following sections.

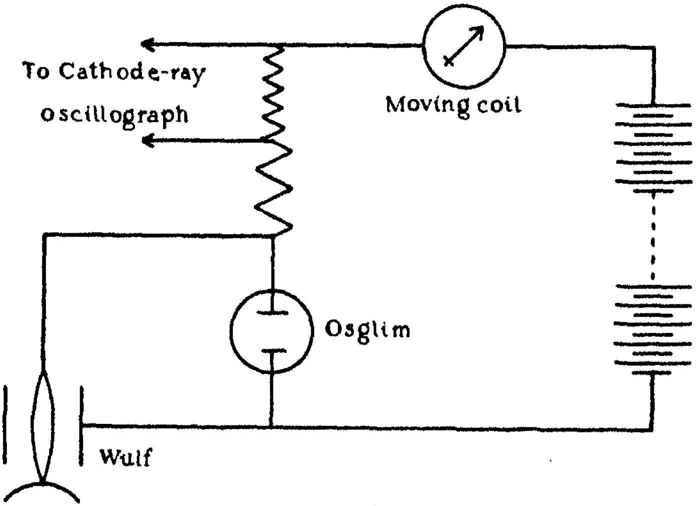


FIG. 2—Circuit for observation of equilibrium.

3—EQUILIBRIUM

The general form is well known, but some particular features need attention.

To clear up a doubt, expressed by Shaxby and Evans (1924) (see p. 296), a cathode-ray oscillograph was included in the circuit to show for certain whether the current was steady, and not the average of an intermittent effect. Those recorded in § 3 were certainly steady. The oscillograph took an appreciable current, of the order of a few microamperes. So, for the smallest currents in the osglim, the oscillograph could not be connected directly across the osglim. Instead, the voltage across the osglim was observed by a Wulf electrometer of resistance > 10^{16} ohms; the current through the osglim was observed by a moving-coil instrument and the oscillograph was put across a series resistance as in fig. 2.

The range of currents was from 0·5 to 7700 μA. The general results are shown in fig. 3, but the smaller currents are plotted on a much magnified scale in fig. 4, Plate 13.

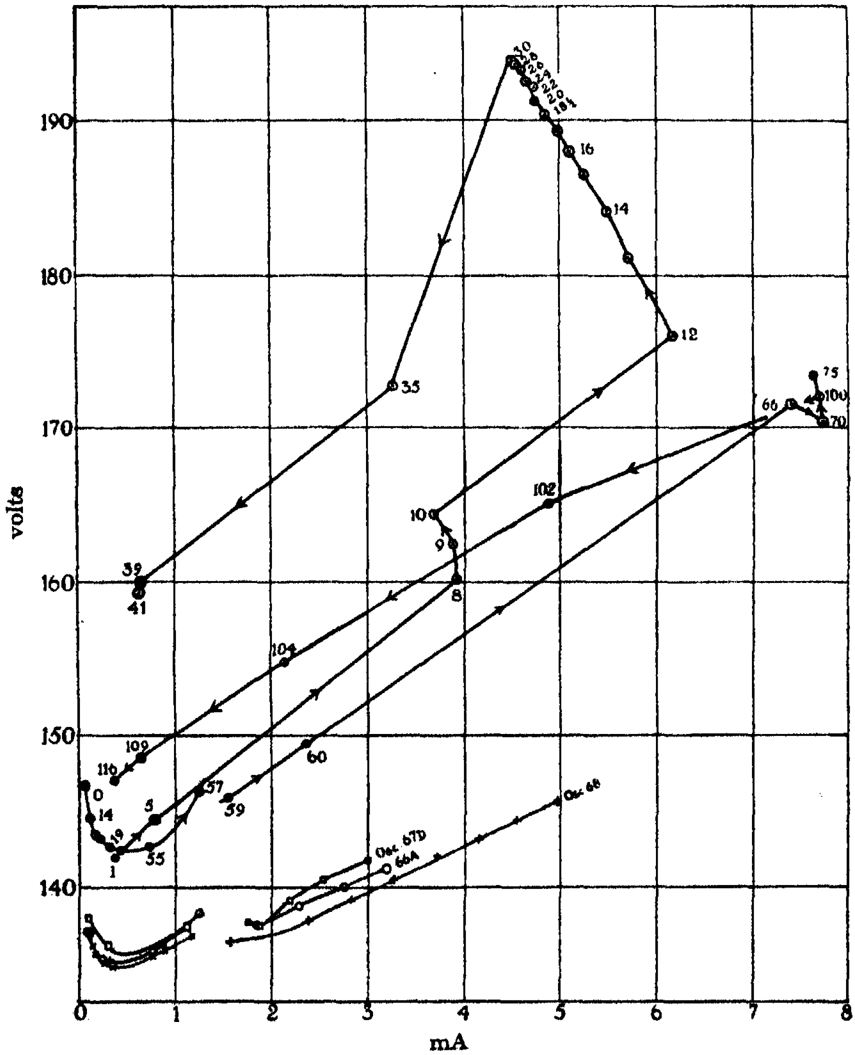


FIG. 3—General view of equilibrium.

The collection of curves seems to have a definite low-voltage limiting curve. For simplicity this will be discussed first.

The surmise of Shaxby and Evans (1924) was confirmed. What Pearson and Anson (1922*b*) had called the upper and lower critical potentials are merely two parts of a continuous equilibrium curve.

There is indeed a discontinuity, but it is in another place, on the large-current side of the minimum voltage at about 1·3 mA. It is to be seen on Oscillogram 68. It was observed more in detail when the osglim was connected in a telephone bridge. For a click was heard in the telephone at the instant when the glow was seen to jump from one side only of the cathode to both sides; and when the current was diminished there was another click as the glow jumped back again. The current at which the jump occurred varied by ± 0·3 mA from day to day, but on the same occasion it was always less when the current was decreasing than when it was increasing. That is to say there is a range extending normally from 1·1 to 1·5 mA in which $\phi_e(u)$ is a two-valued function.

The low-voltage limiting curve is approximately straight for the smallest currents and again for the largest currents, and has the following slope resistances:

For currents between 2500 and 5000 μA, $\rho = +2450$ ohms.

For currents between 0·5 and 50 μA, $\rho = -2{\cdot}8 \times 10^5$ ohms.

There is an effect of temperature. Pearson and Anson (1922*b*) showed that their "lower critical voltage" was increased by 22 V when the temperature of the medium surrounding the osglim was raised from 0 to 100° C. In the present experiments a similar but rather smaller temperature coefficient, namely +0·16 V/1° C., was observed between 10 and 40° C., when the current was such as to make the voltage a minimum, and was allowed to pass for so short a time as not to produce much drift. But usually the electrical history of the osglim is of more consequence than the room temperature.

Drift and recovery must now be considered. The lower curves with crosses in fig. 3 were recorded on an oscillogram (Osc. 68) with brief currents in August 1934 in a room at 17·6° C. Many other observations taken at about this time gave nearly parallel curves within a volt or two of that shown. But the passage of current raised the equilibrium voltage. This is clearly shown on Oscillogram 72, fig. 4, Plate 13, where the numbers placed beside the dots are time in half-minutes from the start. The osglim had been free from current for 2 days. All the experiment was done with as brief currents as possible, except between 5 and 25 on the scale of half-minutes. During that period the current persisted; and a notable drift occurred. Part of the drift persisted when the current was diminished.

Similar but much larger effects are shown in the upper part of fig. 3. The osglim having been free from current for 6 days, observations began on 7 November 1936 at a small current, at a point marked *O* on fig. 3. The numbers beside the graph are time in minutes from the beginning of the run. There was a minimum voltage of 142·4 V at 12°·5. So the minimum had risen 6 V in 2·3 years; certainly not a temperature effect, for that would be 0·8 V in the opposite sense. The current was then increased to 7·5 mA, each eye observation being taken with extreme rapidity and the current immediately switched off. See the points marked 66, 70, 75, 100. On testing again 16 min. later at 0·375 mA the voltage is seen to be 4·5 V higher than it was 97 min. earlier at the same current.

At first sight the loop might be called a hysteresis loop; but that name might be misleading. Because in magnetic hysteresis, time-order alone matters; whereas for the osglim, duration has also a great effect.

After nearly 24 hr. rest, the osglim appeared to have quite recovered, as judged from the first observation; which is marked 1, and is even a little below the minimum of the previous day. The succeeding observations numbered 5, 8, show, however, that the appearance of complete recovery was illusory. The numbers placed beside the points are time from the start in minutes. From 12 to 30 the current was flowing continuously. A huge drift occurred in this period. The voltage rose by 16 V, while the current decreased. The crowding of the numbers towards the top of the line shows how the drift slowed down in a nearly exponential manner, one-half of the ultimate change being attained in about 3 min. This time constant is of the same order as that for heating (see § 2). This resemblance suggests that the drift of the equilibrium voltage may be a consequence of the heating of some solid part of the osglim by the discharge. Now the whole rise of voltage at the point labelled 30 is 30 V above the value for the same current applied briefly. And a rise in external temperature ultimately raised the equilibrium voltage by 0·16 V per 1° C. Therefore, to explain the drift, we have to suppose that some solid part is raised by 30/0·16 = 188° C. This rise cannot be attributed to the inner surface of the bulb; for the power being dissipated was only 0·87 W; whence, from the size of the bulb and the conductivity of glass, it is evident that the inner and outer surfaces would differ in temperature by not more than about 0·1° C.; and the outside was scarcely warm. But it is conceivable that one or both metal electrodes might be raised by 188° C.

After the point labelled 30 the current was switched off and only intermittently passed for the observations marked 35, 39 and 41 min. It is evident from the diagram that recovery at zero current is slower than drift with persistent current. The rise of voltage may perhaps be explained as a consequence of heating, but the fall cannot be explained as an immediate consequence of cooling.

The osglim was now given a week's rest at no current. Then (on 15 November 1936) a set of observations was made with very small brief currents. They are shown in fig. 5 and numbered in minutes as before. The drift was much smaller and in the opposite sense to that at large currents; for the voltage fell as time went on, ending about 1·6 V below that observed on 24 October 1936, before the osglim had been subjected to the large currents.

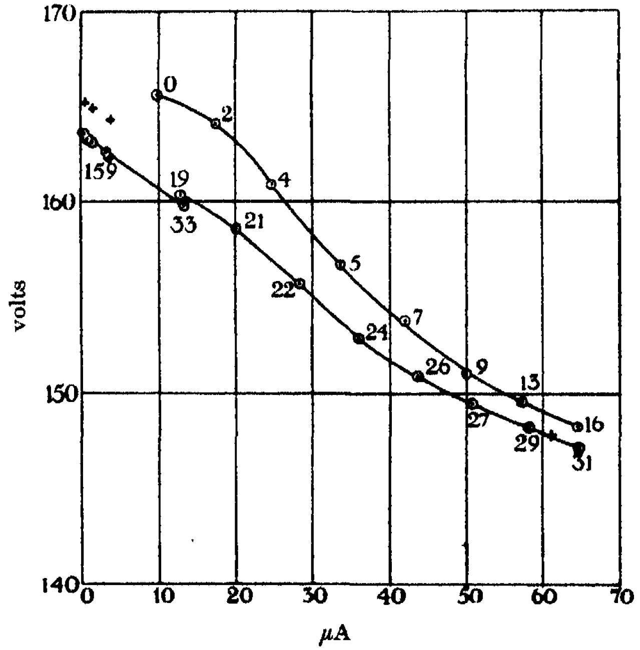


Fig. 5—Equilibrium for small currents. Crosses 24 October 1936. Circles 15 November 1936 with time in mins.

One effect of the passage of current is to make the slope resistance more positive. There are several straight lines on fig. 3 corresponding to a slope resistance of about 4500 ohms, a value of frequent occurrence in the bridge measurements of § 6.

To secure comparable results it is evidently important:

(i) to avoid large currents such as 5 mA,

(ii) to allow for recovery from any unavoidable large currents,

(iii) either (*a*) to observe within a few seconds of the starting of the current or (*b*) to maintain a steady current for about 15 min. before observing,

(iv) and most important, to observe the equilibrium as nearly as possible at the same time as the disturbance with which it is to be compared. On later oscillograms the time interval between photographing the loop and the dots was ordinarily a minute or two, and could hardly be made less.

4—THE INITIATION OF CURRENT

The experiments described in this section were made in the year 1931 on an osglim, labelled L, glowing on the spiral. This lamp resembled osglim P used for all the other experiments. It was found that in the range $2 \times 10^{-6} < u < 3 \times 10^{-5}$ amp., the equilibrium voltage was

$$\phi_e(u) = P + Qu,$$

where $P = 163{\cdot}5$ V, $Q = -3{\cdot}3 \times 10^5$ ohms. It seems reasonable therefore to write $\operatorname*{Limit}_{u \to 0} \phi_e(u) = P = 163{\cdot}5$. But actual observation of $\phi_e(0)$ by increasing ϕ at zero current until a discharge began, gave voltages erratically higher than P; so that

$$\phi_e(0) \geqslant \operatorname*{Limit}_{u \to 0} \phi_e(u).$$

The osglim was connected in a Pearson-Anson circuit and the time of charging was varied from 10 to 0·3 sec. by altering the capacitance. The voltage across the osglim was observed by a Wulf electrometer which had a period of vibration of the order of 0·02 sec. As the speed of charging was increased, the maximum voltage ceased to have any definite value. Discharge began at various voltages lying in a wide range. The lower end of this range remained apparently fixed at about 163 V, but the upper end rose greatly, as if the osglim were waiting for some external event that did not occur more frequently than about 5 times/sec. on the average, and occurred erratically. On covering the osglim with black paper, the flashes became rarer and the highest voltage rose further, sometimes pausing for a visible fraction of a second at the full 220 V given by the battery.

Fig. 6 (Wulf 12) Plate 13 shows the effect when the osglim was entirely surrounded by 0·4 cm. thickness of lead. The time lines are 0·40 sec. apart;

the other scale is 2·02 V/division. It is seen that there are erratic variations of 24 V in the sparking potential.

An explanation, which will be justified in §8 is that the factor $\frac{1}{u}$ in $\frac{b}{u}\frac{du}{dt} = \phi - \phi_e(u)$ causes du/dt to vanish when u vanishes; so that, although the electrical state is unstable above $\operatorname*{Limit}_{u \to 0} \phi_e(u)$, any discharge would begin with indefinite slowness, in the absence of external disturbance. Nothing happens until a stray ray starts the ionization. The osglim then functions rather like a Rutherford-Geiger counter (Valle and Rossi 1935; Valle 1936). When in the lead case it was probably counting cosmic rays, as there was no known radioactive material about.

To test various kinds of light, the osglim was enclosed in a lead box and light from a tungsten lamp was admitted through a hole 0·5 cm. diam. covered with various filters. The active light, which prevents the extreme overvoltage was:

almost perfectly transmitted by quartz, fluorite, or water in layers 0·1–0·3 cm. thick,

partially transmitted by vita glass 2·0 cm. thick; less so by ordinary greenish plate glass 0·94 cm. thick,

entirely stopped by ebonite 0·07 cm. thick, or by red, green or blue gelatine filters.

It was concluded that the active light had a wave-length less than $0{\cdot}33\mu$.

The author has repeatedly looked for an effect of daylight or of tungsten light on the electrical constants of an osglim which was already glowing; but he has found none.

5—SMALL OSCILLATIONS ABOUT EQUILIBRIUM. INTRODUCTION

A glow in the gas was maintained by a steady current from a battery. On this was superposed a much smaller alternating current.

The reason for studying small oscillations before large disturbances is the same for the osglim lamp as it is for the pendulum, or for sound waves: by so doing we avoid, in the theory, terms of degree higher than the first; while in the experiments we avoid the production of harmonic overtones.

Two quite different instruments were used: the kathode ray oscillograph and the alternating current bridge. At frequencies above 5000 c./sec., elliptical oscillograms were obtained; one co-ordinate being voltage, the other being closely related to current. At lower frequencies the ellipse became too narrow to be useful; but A.C. bridges were reliable and convenient. Above 5000 c./sec. the simpler circuit used with the oscillograph was preferable, because it allowed a correction to be made for stray capacitances. In the intervening region near 5000 c./sec. neither method was at its best.

For the convenience of readers who are interested in the psychological analogy, the results are described before the methods are discussed in detail. Indeed a preliminary acquaintance with results is needed before one can judge of the suitability of the methods.

As a preliminary we must make sure that the capacitance and resistance when not glowing do not disturb the effects of glow.

The D.C. insulation resistance between the terminals of the osglim exceeded 3×10^8 ohms.

For alternating potentials measurements were made by a Schering bridge (Table I).

TABLE I

		Equivalent series			Equivalent parallel	
Frequency c./sec.	°C.	Capacitance $\mu\mu F$	Resistance megohms	Power factor	Capacitance $\mu\mu F$	Resistance megohms
560	13	$10{\cdot}_6$	4·1	0·16	$10{\cdot}_4$	180
980	13	$10{\cdot}_5$	$1{\cdot}9_5$	0·13	$10{\cdot}_3$	125
1962	13	$9{\cdot}_4$	$0{\cdot}7_7$	0·09	$9{\cdot}_8$	97
2540	11	$11{\cdot}_1$	$\leqslant 0{\cdot}5_4$	$0{\cdot}09_5$	$11{\cdot}_0$	$\geqslant 60$

Hartshorn (1927), who found that the power factor between the electrodes of an *R*-type valve was 0·20, remarked that "These high-power factors are probably due to absorption effects in the glass."

6—SMALL OSCILLATIONS. RESULTS OF BRIDGE MEASUREMENTS

The bridge measurements are most easily interpreted as giving the slope inductance λ and the slope resistance ρ of the osglim. The way in which λ and ρ vary with frequency and with the superimposed steady current u_0, is made plain by fig. 7. A change of frequency from 50 to 1825 c./sec. causes changes in λ and ρ much larger indeed than those exhibited by good coils, and yet small in comparison with the large ratio of the frequencies.

The superimposed steady current u_0 has a very marked influence on both

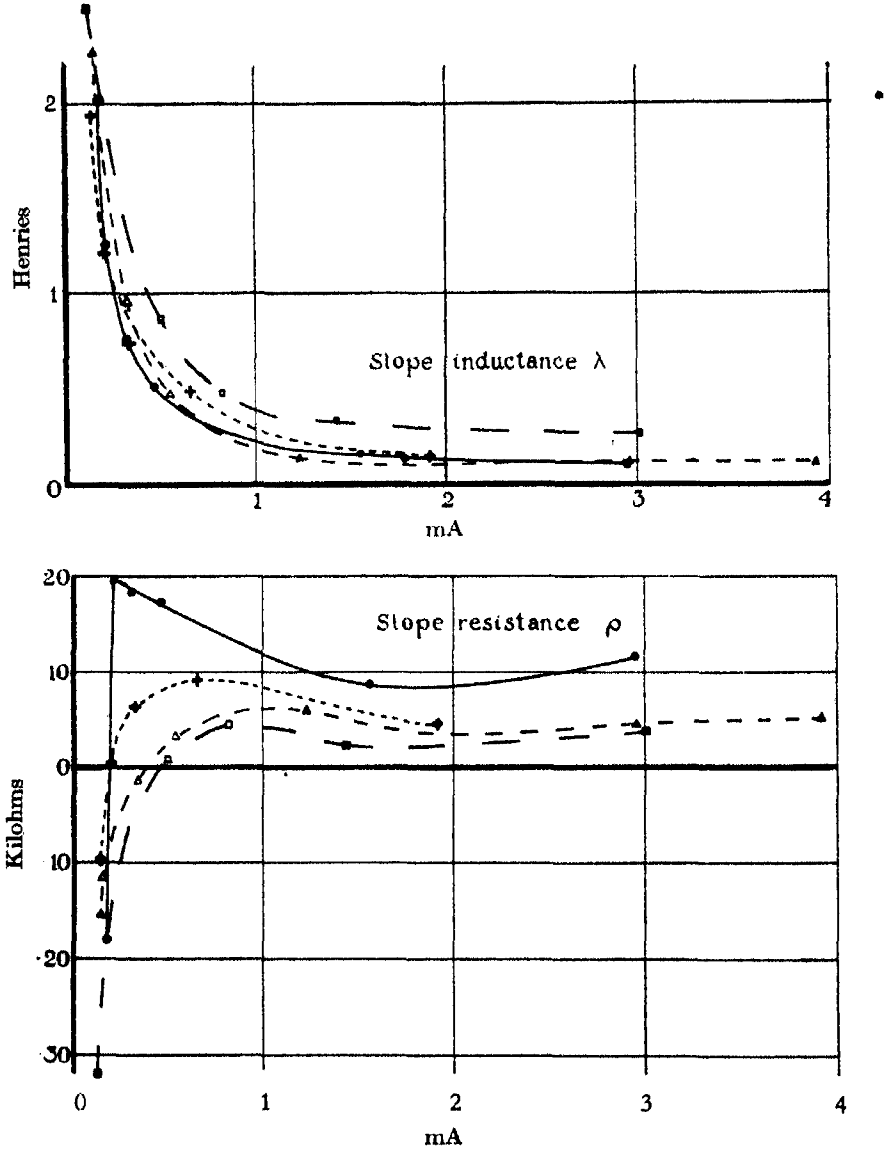


FIG. 7—Bridge measurements of small oscillations.

The words "pip" or "stem" indicate the side of the disk on which the glow occurred.

c./sec.	Glow		c./sec.	Glow	
9805	Pip or both	●	1825	Pip or both	▲
9805	Stem	○	1825	Stem	△
5008	Pip or both	✚	50	Pip or both	■
5008	Stem	+	50	Stem	□

λ and ρ. The slope resistance, as its name suggests, is found to be of the same order as the slope of the equilibrium potential $\phi_e(u)$. That is to say,

$$\rho \doteqdot d\phi_e(u)/du \text{ at low frequencies.}$$

In particular ρ is about 4000 ohms for steady currents lying between 4 and 1 mA, ρ changes sign when u_0 is about 0·3 mA, and ρ has been traced to very negative values such as $-30{,}000$ ohms at $u_0 = 0{\cdot}1$ mA.

On the other hand, the slope inductance λ is of the order of 0·1 H when $u_0 = 3$ mA, and λ steadily increases as u_0 decreases so that λ attains the value 2 H when u_0 is about 0·1 mA. There may be technical applications for this remarkable combination of $-30{,}000$ ohms with 2 H, especially as it has a negligible magnetic field!

At the frequency 1001 c./sec., the equivalent inductance varied inversely as the steady current u_0, the product λu_0 being constant and equal to $2{\cdot}1 \times 10^{-4}$ H-amp. provided the glow only covered part of the stem side of the disk; that is to say for currents up to about 1 mA. For larger currents λu_0 increased somewhat.

The foregoing remarks refer to the values of λ and ρ after the osglim has been glowing for about 15 min.

A cause of variation of λ and ρ is the position of the glow. All the experiments refer to glows near the disk rather than near the spiral, so that the disk was always the kathode. For steady currents u_0 greater than about 1·4 mA the glow entirely covered both sides of the disk. A current of 0·8 mA was sufficient to form a glow all over one or other side of the disk, but not on both sides. A current of 0·3 mA covered only half of either side of the disk. When the glow only partially covered the kathode, any adjustment of the bridge was apt to cause the glow to shift to a different position; and λ and ρ were then usually found to have changed, sometimes by as much as 30 %.

Slow changes. It was observed, in almost every set of bridge measurements, that the slope resistance ρ of the osglim increased with time. When the slope inductance λ of the osglim was held constant, the mean current u_0 also usually increased with time. On the contrary when u_0 was held constant, λ increased with time. In order that results obtained on different occasions may be comparable, it is necessary to conform to a time schedule. Observations taken after $\frac{1}{4}$ hr. glowing were regarded as standard. It is seen from fig. 7 that $\frac{1}{4}$ hr. from the start of a run λu_0 is roughly the same for all u_0. So we may summarize the observations in the statement that ρ and usually also λu_0 increase with time.

At small currents such as $u_0 = 0{\cdot}2$ mA the drift had practically ceased after the osglim had been glowing for 5 min. At large currents, such as $u_0 = 3$ mA, the drift was still considerable after 2 hr. of glowing; as is shown by fig. 8 (Ref. Br 51). So persistent a drift cannot be an immediate effect of the heat produced by the discharge.

Recovery from the changes produced by 15 min. glowing was also very slow in comparison with that for temperature changes. At $u_0 = 0{\cdot}2$ mA it was possible to obtain consistent observations by glowing the osglim for 15 min. every 6 hr., because recovery took place in this time. But at $u_0 = 2$ mA an interval of 6 hr. was not sufficient for recovery. See fig. 9. After the osglim had not been glowed for a week, ρ and λu_0 were at first unusually low, so that it seemed best to reject the first 15 min. glowing, then to leave the osglim extinct for 6 hr. At the next glowing it was in its usual state.

An invisible change prior to lighting—Once when the osglim had not been glowed for a month or more it was connected into a Grover bridge and the battery intended to produce the glow was switched on. During about 2 sec. before the osglim lit, the bridge went progressively more and more out of balance, as was made plain by the swelling sound in the telephone. This showed that there was an invisible change prior to lighting. But that is a rare phenomenon.

Temperature coefficients

To measure these the osglim was kept extinct in transformer oil in a Dewar flask for about 6 hr. It was then lit and 15 min. later the bridge was balanced. It was necessary to go both up and down in temperature as there was considerable hysteresis. The bridge was such that the slope inductance λ was held constant while the mean current u_0 and bridge resistance were adjusted to suit the temperature T. Thus, if suffixes denote what is constant, the bridge measured $(\partial u_0/\partial T)_\lambda$ and $(\partial \rho/\partial T)_\lambda$. Table 11 shows the values found in the neighbourhood of room temperature.

TABLE II

$(\partial u_0/\partial T)_\lambda$ mA/° C.	$(\partial \rho/\partial T)_\lambda$ kilohm/° C.	u_0 at 18° C. mA	Frequency c./sec.	Reference number of observations
$-0{\cdot}002$	$-0{\cdot}6_6$	$0{\cdot}21$	5011	117, 118
$-0{\cdot}008$	$-0{\cdot}3$	$0{\cdot}34$	5008	126, 127, 128
$-0{\cdot}02_8$	$-0{\cdot}2$	$1{\cdot}55$	9800	129, 142, 143, 144
$-0{\cdot}06$	$-0{\cdot}1_1$	$1{\cdot}9$	5008	130, 131, 132, 133

It is seen that the progression of the values of $(\partial u_0/\partial T)_\lambda$ and $(\partial \rho/\partial T)_\lambda$ as functions of u_0 is not much disturbed by a jump in the frequency.

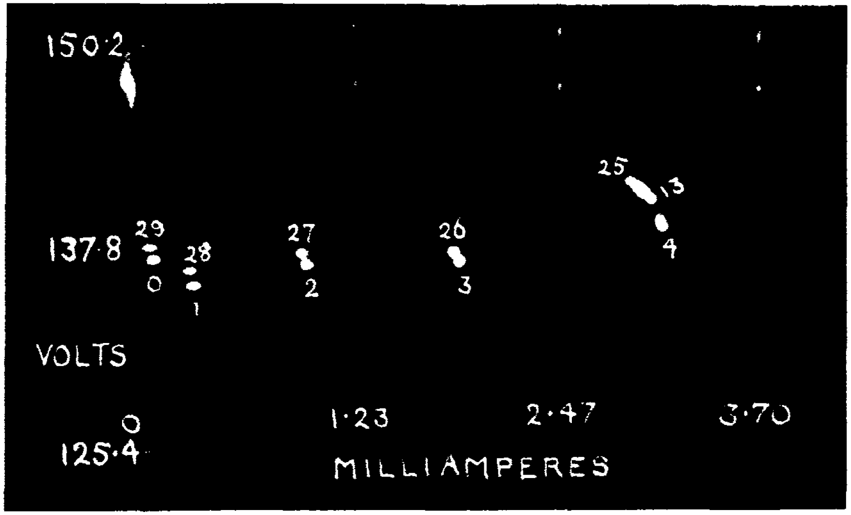


Fig. 4—Displacement of equilibrium by duration of current.

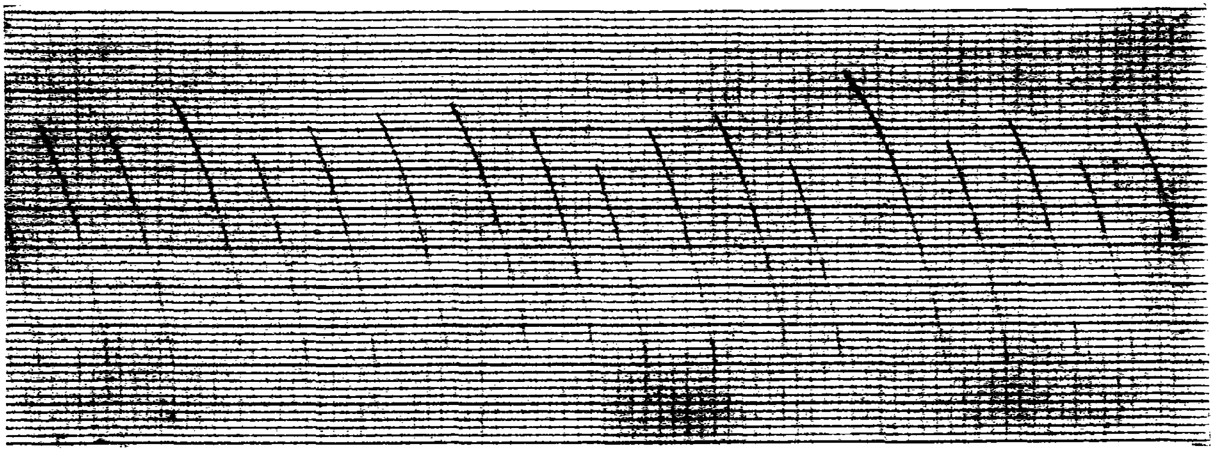


Fig. 6—Erratic sparking potential. Time advances to the left; voltage up the page. The smallest divisions are 0·40 sec. and 2·02 volts.

For many purposes we want to know the coefficients $(\partial\rho/\partial T)_{u_0}$ and $(\partial\lambda/\partial T)_{u_0}$ taken at constant current. Now it can be shown that

$$\left(\frac{\partial\lambda}{\partial T}\right)_{u_0} = -\left(\frac{\partial u_0}{\partial T}\right)_\lambda \cdot \left(\frac{\partial\lambda}{\partial u_0}\right)_T, \quad \left(\frac{\partial\rho}{\partial T}\right)_{u_0} = \left(\frac{\partial\rho}{\partial T}\right)_\lambda - \left(\frac{\partial\rho}{\partial u_0}\right)_T \cdot \left(\frac{\partial u_0}{\partial T}\right)_\lambda,$$

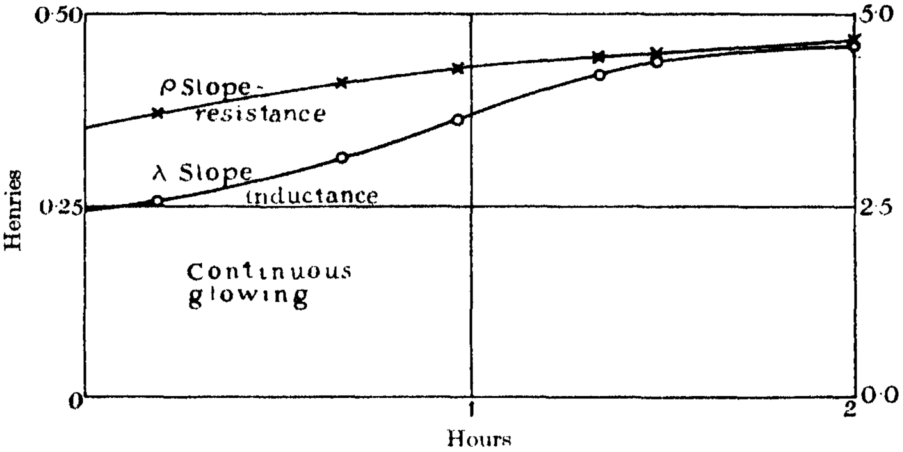


Fig. 8—Continuous glowing by D.C. of $2{\cdot}9_6$ mA, measurements by small superposed A.C. at 50 c./sec., 14·6° C.

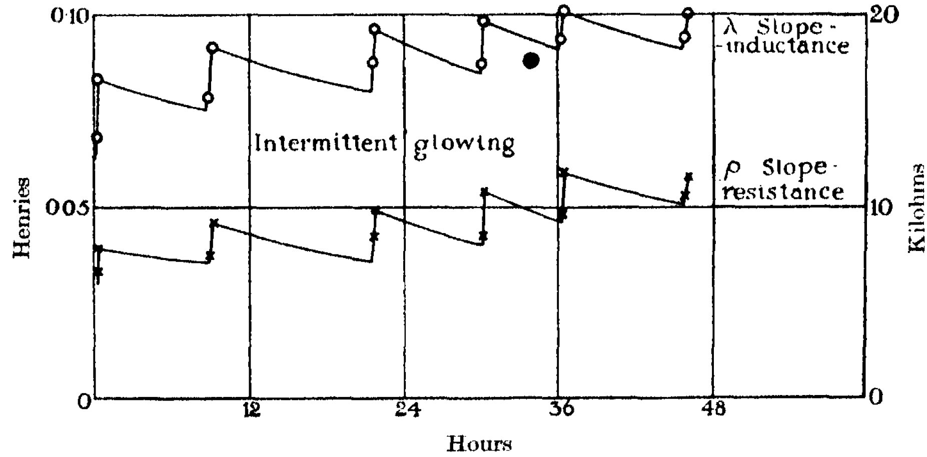


Fig. 9—Intermittent glowing by D.C. of 2·95 mA kept on for 15 min. at a time. Measurements by small superposed A.C. at 9805 c./sec., 17·5° C. Ref. Br 146–151.

in which $(\partial\lambda/\partial u_0)_T$ and $(\partial\rho/\partial u_0)_T$ can be read from the curves in fig. 7. By using these formulae the numbers in Table II were converted to those in Table III. The simplest of these facts is that $\dfrac{T}{\lambda}\left(\dfrac{\partial\lambda}{\partial T}\right)_{u_0}$ is of the order of -4;

so that, at constant mean current, λ varies inversely as the fourth power of the absolute temperature, or nearly so.

TABLE III

u_0 at 18° C. mA	$(\partial\lambda/\partial T)_{u_0}$ H/° K.	$(\partial\rho/\partial T)_{u_0}$ kilohm/° K.	$\frac{T}{\lambda}\left(\frac{\partial T}{\partial\lambda}\right)_{u_0}$ numeric
0·21	$-0·01_1$	$-0·52$	$-2·7$
0·34	$-0·01_1$	$-0·13$	$-4·5$
1·55	$-0·002_3$	$-0·2_5$	$-4·4$
1·9	$-0·003$	$-0·1_1$	$-5·8$
		Mean	$-4·3$

No such simplicity can be expected in the variations of ρ; for as ρ changes sign $(\partial\rho/\partial T)_{u_0}$ remains negative.

7—SMALL OSCILLATIONS. BRIDGE METHODS 50–1825 C./SEC.

There was always a considerable background noise. It was not mains-hum, for it persisted when the house was disconnected from the mains. It resembled the sound of rain, or other sounds caused by a multitude of small impacts. To quieten the background a tunable telephone (by Messrs S. G. Brown, Ltd.) was used. The background is most disturbing at the low audio-frequencies such as 258 c./sec. At this pitch it was found that the bridge could be balanced more accurately when the generator had a ticking grid (Appleton, Watson-Watt and Herd 1926) so that its note sounded intermittently.

The description will be shortened by references to the standard treatises of Hague (1930) and Campbell and Childs (1935). All the bridges were provided with Wagner auxiliary arms.

As the alternating voltage applied to the bridges had to be restricted to a range of about 1 V, the balance was at first insensitive; a fault which was remedied by feeding the detector through a triode amplifier.

The first bridge to be tried had equal ratio resistances and the osglim lamp was balanced against a battery, a resistance and a self-inductance, all in series. By adjusting all these three parts a balance was obtained for both alternating and direct currents. But the interpretation of the results was doubtful, for compact batteries have polarization and bulky batteries have capacitance to earth. So that type of bridge was abandoned.

The omission of the battery left an unbalanced steady electromotive force of about 140 V across the detector, which therefore had to be protected by a condenser; while, to prevent excessive currents in two of the bridge arms, it was desirable to insert a condenser in at least one of them. A bridge of this type was invented by Hay and improved by Hartshorn (Hague 1930, p. 268; Campbell and Childs 1935, p. 242). Hay's bridge was tried for the present purpose with some satisfaction; but the successive adjustment for balance occupied a few minutes during which the condition of the osglim drifted.

Hay's bridge was therefore abandoned in favour of Owen's bridge (Hague 1930, p. 276; Campbell and Childs 1935, p. 244) which could be balanced in less time.

The Owen bridge, on account of its asymmetry, is affected by the impurity of its component parts in a complicated way. Care was taken to use parts of good quality. The condensers had mica or air for dielectric. The larger resistances were either Dubilier metalized 1 W white sticks, contemporaneously standardized for D.C. resistance or else Ferranti anode-feed wire-wound resistances standardized for D.C. The smaller resistances were Cambridge and Paul boxes designed for A.C. bridges. Even thus the residuals of the Owen bridge were too large to be neglected at 50 c./sec.; for at this frequency the problem was to measure a reactance in the presence of a resistance about 100 times larger, this resistance being placed next the osglim to ensure stability. The Owen bridge, in these difficult circumstances, was used only to measure the differences in λ and ρ between the osglim and a coil substituted for it.

Quicker and more reliable than Owen's was a bridge in which balance was attained by adjusting both a resistance and the steady current u_0. The speed of operation depended on the fact that the main ratio arms and Wagner arms were both set beforehand. The arrangement, apart from the special method of balancing and an interchange of source and detector, was that known as Grover's (Hague 1930, p. 276). The coils with which the osglim was compared were standardized at the same frequencies either by Dye's shunt bridge (Hague 1930, p. 306; Campbell and Childs 1935, p. 253 correcting the formulae given by these authors) or by Campbell's unequal ratio bridge (Hague 1930, p. 298; Campbell and Childs 1935, p. 233) or at the higher frequencies by Gruneisen and Giebe's resonance bridge (Hague 1930, p. 244; Campbell and Childs 1935, p. 271), with a correction for stray capacitance to earth. For example one of the larger coils (R36/2) varied with frequency in the manner shown in Table IV.

Bridge methods at 9800 *and* 5008 *c./sec.*

The difficulties usual in bridges at such frequencies and caused by stray capacitances were here complicated by drift and increased by the large value of the impedance to be measured. Whereas in the admirable bridge of Dye and Jones (1933) the arms were of the order of 300 ohms, here they had to be 30,000. The author can hear 5000 c./sec. but not 9800 c./sec. A small detector could not be used, because the only kinds available were the vacuo-junction and the metal rectifier, and these, being unselective, were disturbed by the background "rain-noise" of the osglim. The bulkier heterodyne detector was made as symmetrical as possible by using two tetrodes in push-pull connexion and by earthing their midpoint. The generator and its leading wires were enclosed in earthed screens. Separate screens enclosed the detector, the ammeter and the battery for the superposed direct current. But it was found to be best not to earth any screen except that of the generator and its leads.

TABLE IV

°C.	c./sec.	Ohms	H
9·8	0	4371	—
17·1	0	4510	—
14·8	1000	4462	$1{\cdot}192_7$
13·6	5011	4600	$1{\cdot}21_1$
13·7	9813	5030	$1{\cdot}25_9$

The arms of the bridge were not screened, but were well separated from each other and from the table and were connected by thin wire, 0·0193 cm. diameter. An adjustable resistance of the order of 10^4 ohms had to be in series with the bridge in order to regulate the direct current. Experience showed that it was important to divide this resistance into two equal rheostats, one on each side of the bridge. It was found to be essential to use a bridge with equal arms in order to be able to make reversals. The arms were Dubilier spirohms, nominally 30,000 ohms, and capable of carrying 15 mA. Two kinds of reversal were made: the generator wires were interchanged, and the ratio arms together with the detector were interchanged *en bloc*. The arithmetic mean of the four observations was regarded as significant.

8—DIFFERENTIAL EQUATIONS FOR SMALL OSCILLATIONS AT AUDIO-FREQUENCIES

The sort of formula here required is one which shall describe the salient features in an easily understandable manner; because the irregularities depending on the irregular shape of the electrodes are not interesting; and because, for later application in psychology, an easy approximate formula will be preferred to one that is precise but cumbrous.

From the definition of the slope resistance ρ and slope inductance λ we have,

$$\phi-\phi_0=\lambda\frac{d(u-u_0)}{dt}+\rho(u-u_0), \tag{1}$$

where u_0 is the superposed direct current.

At the lowest frequency, 50 c./sec., the results of bridge measurements show that

$$\rho_{50}=\frac{d\phi_e(u)}{du}\text{ approximately,} \tag{2}$$

also that

$$\lambda_{50}=a+\frac{b}{u_0}, \tag{3}$$

where, corrected to 18° C.,

$$\left.\begin{aligned} a&=0{\cdot}14\,\text{H},\\ b&=2{\cdot}8\times10^{-4}\,\text{H-amp.}\end{aligned}\right\} \tag{4}$$

Inserting (2) and (3) in (1) and transposing we have

$$\phi-\phi_0-(u-u_0)\frac{d\phi_e}{du_0}=\left(a+\frac{b}{u_0}\right)\frac{d(u-u_0)}{dt}. \tag{5}$$

But by Taylor's series,

$$\phi_e(u)=\phi_0+(u-u_0)\frac{d\phi_e}{du_0}+\frac{(u-u_0)^2}{2}\frac{d^2\phi_e}{du_0^2}+\ldots, \tag{6}$$

and because the oscillations are small the series will be cut off after the linear term. Insertion of (6) in (5) gives

$$\phi-\phi_e(u)=\left(a+\frac{b}{u_0}\right)\frac{d(u-u_0)}{dt}. \tag{7}$$

For small oscillations the term b/u_0 is indistinguishable from b/u; also $du_0/dt=0$. Accordingly (7) takes the simple form

$$\phi-\phi_e(u)=\left(a+\frac{b}{u}\right)\frac{du}{dt}, \tag{8}$$

at low frequencies, where a and b are constants given by (4). So we have come back to something rather like Braunbek's theoretical result mentioned in the introduction p. 297, eqn. (5). At small currents a is negligible relative to b/u.

As the frequency increases, λ decreases and ρ increases. These changes, as shown on fig. 7 are not great, considering the wide range of frequency from 50 to 9800 c./sec. Equation (8) does not need much alteration at 1000 c./sec.

The peculiar forms of the resistance curves at the higher frequencies may be obtained from the resistance curve at 50 c./sec. by adding to its ordinate a fixed multiple of the ordinate of one of the inductance curves. In fact we have

$$\rho_{9800} = \rho_{50} + 2 \times 10^4 \lambda_{50} \tag{9}$$

$$= \frac{d\phi_e(u)}{du} + 2 \times 10^4 \left(a + \frac{b}{u_0}\right), \tag{10}$$

the units being ohms, henries, volts and amperes.

Whether or not there is a corresponding cross-connexion for the decrease in λ with increasing frequency, is not so evident from an inspection of fig. 7. The various λ curves certainly do cross one another near the currents at which ρ changes sign. The formula,

$$\lambda_{9800} = \lambda_{50} - \frac{\rho_{5000}}{3 \times 10^4}, \tag{11}$$

describes the main features. But, on the same plan as (9),

$$\rho_{5000} = \rho_{50} + 8 \times 10^3 \lambda_{50}. \tag{12}$$

Elimination of ρ_{5000} between (11) and (12) yields

$$\lambda_{9800} = 0{\cdot}73\left(a + \frac{b}{u_0}\right) - \frac{10^{-4}}{3}\frac{d\phi_e(u)}{du}. \tag{13}$$

The insertion of (10) and (13) into (1) yields in the same manner as before,

$$\phi - \phi_e(u) + \frac{10^{-4}}{3}\frac{d\phi_e(u)}{dt} = 0{\cdot}73\left(a + \frac{b}{u_0}\right)\frac{du}{dt} + 2 \times 10^4 \left(a + \frac{b}{u_0}\right)(u - u_0), \tag{14}$$

for small oscillations at 9800 c./sec. It will be shown later that equation (14) is suggestive in regard to the forms of very large oscillations at that frequency.

The observations shown on fig. 7 were made at various times during 18 months. They cover a wide range. But they are not the best way of measuring changes due to changing frequency; for drift may have intervened.

9—SIMULTANEOUS OBSERVATIONS AT DIFFERENT FREQUENCIES

For the elimination of drift, observations at frequencies A, B, C were made in time sequences such as $ABCBA$, $BCACB$ or $CABAC$ spread over about ½ hr. The time of each bridge setting was recorded to the nearest minute. The observed λ and ρ were corrected to a standard time, 15 min. from the start of glowing. The osglim was then given a rest, usually of a day, or half a day. The bridge connexions were reversed and the observations repeated in a different one of the aforesaid time sequences. For the elimination both of stray fields and of the inequalities in the ratio-arms, four sets of readings were taken by reversing separately the generator connexions and the detector connexions. The results, which were obtained in 1936, are shown in Table V.

The fitting of formulae to these observations will be delayed until they can be combined with corresponding oscillograph measurements at higher frequencies.

TABLE V

u_0 mA	λ H	ρ kilohms	Frequency c./sec.	°C.	Ref.
0·62	0·47	6·0	3530	16·7	Br 185–188
	0·49	4·6	1826		
0·62	$0\cdot6_3$	6·6	1826	18·5	Br 202–209
	$0\cdot_8$	5·	258		
	1·	$4\cdot_9$	50		
2·95	$0\cdot17_4$	5·5	1825	$17\cdot_7$	Br 197–201
	$0\cdot2_1$	5·	258		
	$0\cdot3_1$	$5\cdot_0$	50		

Observations in the same horizontal compartment of the table have been corrected to the same time.

This research was made possible by two benefactions. The late Mr. William Bow gifted for the use of the Principal of Paisley Technical College a dwelling house large enough to contain a laboratory; and the Government Grant Committee of the Royal Society provided a considerable part of the apparatus. Mr. Albert Campbell has kindly advised me on a number of problems concerning the bridge measurements.

REFERENCES

Appleton, E. V., Watson-Watt, R. A. and Herd, J. F. 1926 *Proc. Roy. Soc.* A, **111**, 615.

Braunbek, W. 1926 *Z. Phys.* **39**, 8.

Campbell, A. and Childs, E. C. 1935 "Measurement of Inductance, Capacitance and Frequency." London: Macmillan.

Dye, D. W. and Jones, I. 1933 *J. Instn Elect. Engrs*, **72**, 169–81.
Garnett, J. C. M. 1919*a* *Proc. Roy. Soc.* A, **96**, 91.
— 1919*b*. *Brit. J. Psychol.* May 1919.
Griffith and Taylor 1925. *See* Taylor 1925.
Hague, B. 1930 "Alternating current bridge measurements", 2nd ed. London: Pitman.
Hartshorn, L. 1927 *Proc. Phys. Soc.* **39**, 108–23.
Leyshon, W. A. 1930 *Proc. Phys. Soc.* **42**, 157–69.
— 1932 *Proc. Phys. Soc.* **44**, 171–89.
McDougall, W. 1901 *Brain*, **24**, 577.
Pearson, S. O. and Anson, H. St. G. 1922*a* *Proc. Phys. Soc.* **34**, 175–6.
— — 1922*b*. *Proc. Phys. Soc.* **34**, 204–14.
Pol, B. van der 1926 *Phil. Mag.* **2**, 978–92.
Pol, B. van der and Mark, J. van der 1928 *Phil. Mag.* **6**, 763–75, and exhibit at a subsequent Soirée of the Royal Society.
Richardson, L. F. 1930 *Psychol. Rev.* **37**, 214–27.
— 1932 *Proc. Phys. Soc.* **44**, 189.
Shaxby, J. H. and Evans, J. C. 1924 *Proc. Phys. Soc.* **36**, 253–61.
Taylor, G. I. 1925 "The Mechanical Properties of Fluids", Chap. VII. London; Blackie.
Valle, G. 1936 *Sci. Abstr.* A, 167.
Valle, G. and Rossi, B. 1935 *Sci. Abstr.* A, 2070.

The Behaviour of an Osglim Lamp
II—Oscillograph Methods, Steady Motion and Relaxational Oscillations

By Lewis F. Richardson, F.R.S.

(*Received* 1 *June* 1937)

[Plates 14–16]

10—Introduction

The large oscillations have been studied by means of a cathode-ray oscillograph. The technique is necessarily somewhat elaborate and is described in § 11, which can be skipped by any reader who is interested only in the results. The aim of the investigation was to find an equation to describe the behaviour of the osglim under all non-destructive circumstances. From the study of small oscillations in Part I we bring the fact that

$$\phi - \phi_e(u) = \left(a + \frac{b}{u}\right)\frac{du}{dt}$$

describes small oscillations at 50 c./sec., where $a = 0{\cdot}16$ H, $b = 2{\cdot}9 \times 10^{-4}$ H-amp., while at 1825 c./sec. a becomes negligible. This equation has an intelligible meaning even for the largest oscillations. It will now be shown to provide a good approximate description of a number of interesting phenomena at frequencies up to about 1000 c./sec. The investigation of higher frequencies is in progress, but is not reported yet.

11—OSCILLOGRAPH METHODS FOR OSCILLATIONS BOTH SMALL AND LARGE

The osglim lamp was connected in series with a resistance, and the voltage across the lamp was led to one pair of deflector plates of the oscillograph, the voltage across the resistance to the other pair. Time was recorded on the same trace by unfocusing the spot at regular intervals by applying an oscillating potential to the Wehnhelt cylinder, in the manner described in detail elsewhere (Richardson 1933, 1935).

The resistance which was thus used to measure the current had to be very free from inductance and self-capacitance. A Ferranti 9397 ohm wound in grooves was first employed, then Dubilier metallized sticks of about 10,000 ohm, 1 W and 3 W. The last, being run far below its rated wattage, was particularly free from impurity, but its D.C. resistance drifted by a few parts in a thousand from day to day. [*Note added in proof* 14 *Aug.* 1937—A more constant 10^4 ohm, and with sufficiently small residuals that can be estimated from geometry, is a coil of 860 turns of wire, $0{\cdot}005_5$ cm. diameter wound on a strip $1{\cdot}3 \times 0{\cdot}15 \times 10{\cdot}1$ cm.^{3}; the strip being bent into a **C** shape.]

The interesting range was usually between 120 and 170 V. To explore this with sufficient detail it was necessary to throw the zero of voltage off the screen. Leyshon (1930) effected this by a potential divider, others have used permanent magnets to deflect the cathode rays. The author tried both these methods. The potential divider led to troublesome computations; the magnets made it impossible to focus the spot sharply on all parts of the trace. More satisfactory results were obtained by interposing batteries between the oscillograph and the potential to be measured, as shown in fig. 1. A great advantage of having batteries in this situation is that they can be used to superpose on the oscillogram, by successive exposures of the same photographic plate, a lattice of standard voltages. This was done by means of the switches shown in fig. 1 and by means of wanderplugs. Immediately after taking the composite photograph, these "lattice batteries" were measured by a substandard voltmeter by Elliot, with a correction obtained by observing the change in the reading of an electrostatic voltmeter (Wulf) when the moving-coil voltmeter (Elliot) was switched in parallel. This correction was as if the "Ever-ready" battery had an internal resistance of about ¾ ohm/V; but the effect is probably due to polarization as well as to resistance.

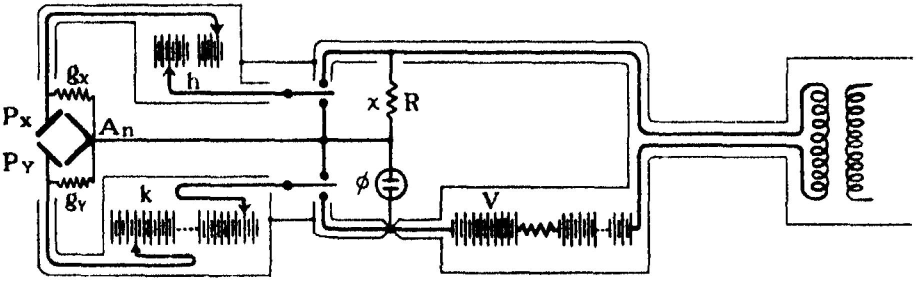


FIG. 1—Measuring circuits and driving coil.

The question arose as to whether the lattice battery gave a sufficiently free passage to high-frequency currents. Two old "Ever-ready" batteries of 120 nominal V were therefore put in opposition and balanced in an equal-ratio bridge against a resistance *R* in parallel with a condenser *C*. The following results were obtained:

Frequency c./sec.	*R* ohms	*C* μF
1000	59	0·90
1826	53	0·40
c. 3530	49	0·17

A pair of new batteries had less *R* and greater *C*. These values are satisfactory, as the stray capacitance to earth at the oscillograph end of the battery did not exceed 50μμF. Another make of battery, having a much larger internal resistance, caused trouble.

The voltmeter, which is compensated for temperature, was in turn calibrated in the years 1931, 1934, 1936 against a Weston cell and a resistance box, both with N.P.L. certificates; and was found to drift by 1·3 V or less in the 5 years. In this way the voltages on the oscillograms were recorded with an accuracy of about ½ V at 150 V; while the differences between successive lattice lines were constant to 1/10 V.

For economy four photographs, distinguished as *A*, *B*, *C*, *D*, could be taken on each plate. The plates were numbered serially. This notation appears in print for the convenience of the author, who hopes that the reader will pardon its complexity.

Various means for exciting the neon lamp have been used. One of the most convenient was a battery giving about 150–180 V in series with a coil in which an alternating voltage could be induced, as shown in fig. 1. In other experiments a different circuit containing a condenser, resistance and battery arranged as in fig. 2 was used to give flashes of Anson-Pearson type. But, as the chief problem is to find a description of the behaviour of the osglim lamp valid for *any* external connexions, it would merely distract attention from the real difficulties to consider these somewhat irrelevant circumstances at length.

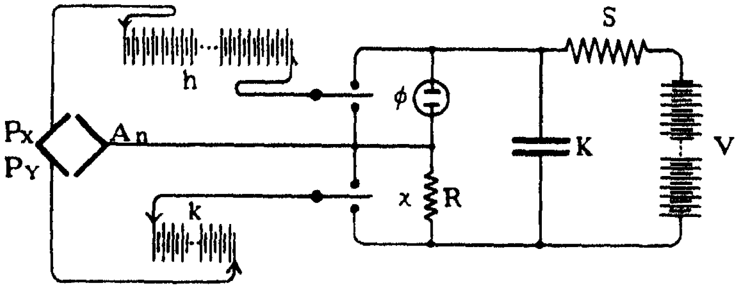


FIG. 2—Circuit for the study of relaxational oscillations.

A small error is due to the current taken by the oscillograph. Resistors, of about 1–4 megohms each, were placed in parallel with each pair of deflector plates up to Osc. 38, but this leak was sufficient to disturb some experiments on Anson-Pearson flashing. So the resistors were removed and not used from Osc. 39 to Osc. 100 except when strays were being measured.

The ionization current between the deflector plates of the oscillograph is well known. In one quadrant of the fluorescent screen this leak was less than 10^{-6} amp. and is negligible. Some oscillograms were therefore taken in this quadrant (Osc. 41).

The static ionization currents were measured by a galvanometer and were recorded as functions of V_x and V_y, the potentials by which the plates P_x, P_y respectively exceed A_n (Table I).

To be sure that the corrections were put in rightly it was unfortunately necessary to lay down the following rather tedious conventions.

List of Symbols with Conventions of Signs

A_n is the anode and the two deflector plates permanently joined to it.

P_x, P_y are the plates which deflect the cathode rays x-ward and y-ward.

V_x, V_y are the voltages between P_x, P_y respectively and A_n. The signs of V_x, V_y are those of P_x, P_y relative to A_n.

h, k are the voltages of the lattice batteries; h for x, k for y; their signs being those of the ends joined directly to the oscillograph.

l is the leak into the plate A_n from plates P_x and P_y jointly.

ϕ is the voltage across the osglim, its sign being that of the end of the osglim attached to the anode.

R is the current-measuring resistance.

χ is the voltage across R. The sign of χ is that of the end attached to the lattice battery, so that $(\chi+\phi)$ is the voltage applied to drive current through R and osglim in series.

Applying Kirchhoff's law to the two circuits, $\chi+h-V_x = 0$. So $\chi = V_x-h$. If $h = 0$, $\chi = V_x$ simply. If $\chi = 0$, $V_x = h$, as happens when the lattice is being made.

Also $\phi-k+V_y = 0$. So $\phi = k-V_y$. If $\phi = 0$, $V_y = k$, as when the lattice is being made.

TABLE I—LEAK FROM PLATES P_x AND P_y TO ANODE ACROSS OSCILLOGRAPH WHEN ANODE VOLTS $= 256{\cdot}_3$

Filament current $= 0{\cdot}9_3$ amp.

The numbers in the body of the table are milliamperes

	V_x in volts						
V_y volts	−37·1	−24·7	−12·4	0	12·4	24·7	37·1
−37·1	—	—	0·000	0·003	0·006	—	—
−24·7	—	0·000	0·000	0·003	0·007	0·009	—
−12·4	0·000	0·000	0·000	0·003	0·007	0·009	0·011
0	0·004	0·004	0·004	0·007	0·010	0·012	0·013
12·4	0·011	0·011	0·011	0·012	0·016	0·017	0·018
24·7	—	0·013	0·014	0·015	0·017	0·018	—
37·1	—	—	0·015	0·018	0·018	—	—

Stray Currents—Conductive strays, except through the gas inside the oscillograph, were negligible. Capacitative strays were negligible at 50 c./sec., but vitally important at 50,000 c./sec. The treatment of strays only gradually arrived at thoroughness. The latest method will be described. Earlier *approximations are briefly mentioned, with the remark that they* will be reconsidered in the light of present knowledge before the earlier experiments are offered for publication. From Osc. 56 to Osc. 99 metallic screens were placed around all the bulky apparatus, except those articles joined directly to the anode of the oscillograph. The screens were connected to one another and to the outer end of the current-measuring resistance R, that is to the end not directly joined to the osglim. The strays were at that time regarded as equivalent to a capacitance C and a resistance R_c both in parallel with R; and C and R_c were measured. Accordingly the current u in the osglim was then taken to be $u = \frac{\chi}{R^*} + C\frac{d\chi}{dt}$, where $R^* = \frac{RR_c}{R+R_c}$.

This procedure certainly corrected most of the error.

Later, at Osc. 100, it was realized that, because the various contents of the screen have considerable alternating voltages between them, the stray Faraday tubes cannot be represented by a single lumped capacitance C, but require two representatives functioning separately. Accordingly, the strays are now assumed to be equivalent to a capacitance E_1 and a conductance E_0 both in parallel with R; and similarly F_1 and F_0 in parallel with the osglim. The current u in the osglim is therefore

$$u = \left(\frac{1}{R} + E_0 + E_1\frac{d}{dt}\right)\chi - \left(F_0 + F_1\frac{d}{dt}\right)\phi, \tag{1}$$

and experiments were made to determine E_0, E_1, F_0, F_1. The leaks E_0, F_0 consist partly of the conductances g_x, g_y of resistors of the order of a megohm intentionally placed between the anode and deflector plates, partly of dissipation in the glass stem of the osglim (see § 5, Table I) and partly also of unavoidable leak g through the conducting gas in the oscillograph. We have to assume that g is the same for both pairs of plates so that

$$E_0 = g_x + g, \tag{2}$$

$$F_0 = g_y + g. \tag{3}$$

Theory indicates that to make $E_1 > F_1$ the screen should be connected to the outer end of R, as it was from Osc. 56 to Osc. 99 inclusive; and that to make $F_1 > E_1$ the screen should be connected to the outer end of the test body (here the osglim) as the screen was from Osc. 100 onward. Experiments confirmed this expectation. Unfortunately, it appears to be impossible to make the ratio $E_1 : F_1$ either zero or infinite.

To determine E_1 and F_1 at any frequency, experiments were made in pairs. For both R was made infinite and the direct voltage was reduced until the osglim did not glow, so that $u = 0$. Accordingly from (1), (2), (3),

$$\left(g_x + g + E_1\frac{d}{dt}\right)\chi = \left(g_y + g + F_1\frac{d}{dt}\right)\phi.$$

In the first experiment an air condenser was put in parallel with the lesser of E_1, F_1 and adjusted to C' until the oscillogram was a piece of the straight line $\chi = k\phi$, the conductances g_x, g_y being adjusted if necessary to prevent the line becoming an ellipse. Thus we have

$$F_1 = (E_1 + C')\,k = kE_1 + C, \tag{5}$$

where C is a known capacitance, commonly of the order of 10^{-10} F.

In the second experiment the air condenser was removed and the elliptical oscillogram was photographed. By measurement of the ellipse $\hat{\chi}/\hat{\phi}$ was found, also the phase angle δ by which $\vec{\phi}$ leads on $\vec{\chi}$. The circumflex denotes the amplitude and the arrow denotes the vector on the phase-and-amplitude diagram. From (4) it can be shown that if $g_x = g_y$ then

$$\delta \text{ is positive if } E_1 > F_1, \quad \text{but negative if } E_1 < F_1. \tag{6}$$

Thus the otherwise ambiguous sign of δ was fixed. Let $i = \sqrt{-1}$ so that

$$\vec{\chi} = (\hat{\chi}/\hat{\phi})\,(\cos\delta - i\sin\delta)\,\vec{\phi}. \tag{7}$$

Also from (4) and (5) on writing ω for the pulsatance we have

$$(g_x + g + i\omega E_1)\,\vec{\chi} = \{g_y + g + i\omega(kE_1 + C)\}\,\vec{\phi}. \tag{8}$$

When $\vec{\phi}$ has been eliminated between (7) and (8) and the resulting pair of equations solved for E_1 and g it is found that

$$g = \frac{-g_x(\hat{\chi}/\hat{\phi})^2 + (\hat{\chi}/\hat{\phi})\,\{\cos\delta(g_y + kg_x) - \omega C\sin\delta\} - kg_y}{(\hat{\chi}/\hat{\phi})^2 - (k+1)\,(\hat{\chi}/\hat{\phi})\cos\delta + k}, \tag{9}$$

$$\omega E_1 = \frac{(\hat{\chi}/\hat{\phi})\,\{\omega C\cos\delta + (g_y - g_x)\sin\delta\} - \omega C}{(\hat{\chi}/\hat{\phi})^2 - (k+1)\,(\hat{\chi}/\hat{\phi})\cos\delta + k}, \tag{10}$$

from which g and E_1 can be determined. But if g is negligible or if $\sin\delta = 0$, it is much easier to determine E_1 from the equivalent equation,

$$\omega E_1 = \frac{\omega C + (\hat{\chi}/\hat{\phi})\,(g_x + g)\sin\delta}{(\hat{\chi}/\hat{\phi})\cos\delta - k}. \tag{11}$$

Results when the screen was connected to the outer end of the osglim are shown in Table II.

At 50,000 c./sec. F_1 is vitally important, E_1 is worth taking into account, but g is almost negligible. The variation of g with frequency is understandable when it is remembered that g is the conductance of a gas. Here F_1 includes the equivalent parallel capacitance of the extinct osglim and g may be disturbed by the neglected dissipation in the glass stem at 50 kc./sec. See Part I, Table I.

TABLE II

Frequency kc./sec.	F_1 $\mu\mu$F	E_1 $\mu\mu$F	g micromhs	Osc. number
1·82	—	—	0·04	114 A and B
3·54	124	16	0·25	113
5·011	135	22	0·04	114 C and D
9·817	133	17	0·13	109 A and B
50·0	136	21	−0·63	106 A and D
50·0	130	21	−0·66	108 B and C
Means	132	19		

12—STEADY MOTION NEAR TO EQUILIBRIUM

Osc. 48 (fig. 3, Plate 14) shows one of the two discharges which occurred in every cycle when the osglim was connected with the lighting mains of 200 V, 50 c./sec. The voltage rose at zero current to 177 V. The sudden breakdown has been too rapid to affect the photographic plate. It is followed by a long loop formed of two traces so close together as to be only just discernibly separate. Osc. 50 (fig. 4, Plate 14) shows the same oscillogram when the spot of light was put out of focus $9{\cdot}8_3 \times 10^3$ times/sec. When the frequency of the focus had been slightly decreased, the time marks were seen to move slowly up to 177 V, to jump across the gap, and then to move slowly in two processions going opposite ways along the nearly coincident traces. Thus it was established that the trace in which the current was increasing was the one at the higher voltage. Let it now be assumed that the equilibrium voltage $\phi_e(u)$ lies between the two close traces in such a way as to make

$$\phi - \phi_e(u) = \lambda \frac{du}{dt} \equiv \lambda \dot{u}, \text{ say,}$$

where λ is a slope inductance like that discussed in the bridge measurements but of unknown magnitude. Let suffixes 1 and 2 refer to the ascending and descending traces. Then by elimination of ϕ_e we have

$$\lambda = (\phi_1 - \phi_2)/(\dot{u}_1 - \dot{u}_2).$$

The photographic plate was measured by a microscope. At $\chi = 24{\cdot}5$ V the separation of the traces gave $\phi_1 - \phi_2 = 0{\cdot}9_1$ V. In this neighbourhood the spacing of the time marks gave $\dot{\chi}_1 = 2{\cdot}8_5 \times 10^4$ V sec.$^{-1}$, $\dot{\chi}_2 = -2{\cdot}3_5 \times 10^4$ V sec.$^{-1}$. At these comparatively slow speeds the stray current $E_1 \dot{\chi} - F_1 \dot{\phi}$ is entirely negligible. The current-measuring resistance was 9397 ohms. Therefore $\dot{i}_1 = 3{\cdot}0_3$ amp. sec.$^{-1}$, $\dot{i}_2 = -2{\cdot}5_0$ amp. sec.$^{-1}$ and $\lambda = 0{\cdot}16$ H at 2·61 mA.

Although the interval $|\phi_1 - \phi_2|$ was too small to be measured accurately, yet the experiment is of interest because both the circumstances of the osglim and the method of measurement differ so widely from those of the small oscillations recorded in Part I. The result, $\lambda = 0{\cdot}16$ H, falls among the bridge measurements for similar currents.

13—RELAXATIONAL OSCILLATIONS

These are worthy of special attention because of the many applications of the flashing circuit described by Pearson and Anson (1922).

13·1—*Oscillograms*—The circuit used in the present experiments is shown in fig. 2. A series of oscillograms (Osc. 39–47) were photographed with $S = 6{\cdot}3_2 \times 10^5$ ohms, $R = 9397$ ohms, $V = 183{\cdot}_5$ V, but various capacitances K.

The conductive leaks E_0, F_0 between the oscillograph plates were made small by decreasing from $0{\cdot}9_3$ to 0·8 amp. the current that heated the filament. For the much-measured Osc. 41 (fig. 7, Plate 16) the leaks were quite abolished by moving the oscillogram into that quadrant of the screen where the polarity makes them zero.

Osc. 47B (fig. 5, Plate 15) shows a loop and equilibrium dots. The voltage during the flash fell to 6 V below the equilibrium value. Osc. 43 (fig. 6, Plate 15) shows, superposed on one another, five loops made with different values of K, as recorded among others in Table III, together with observed frequencies n, and the ratio ν_0/n, where $\nu_0 = \frac{1}{2\pi}\sqrt{\left\{\frac{V-P}{SKb}\right\}}$.

At the largest of these capacitances the breakdown of the sparkgap was so sudden that it does not show in the photograph; and the frequency was too irregular to be measured by a stroboscope. As K was decreased the frequency increased and became regular. Also the voltage penetrated more and more below ϕ_e. It is just as if the phenomenon had inertia, which carried the voltage across its equilibrium value.

At the lower frequencies the fall of current proceeds much more slowly

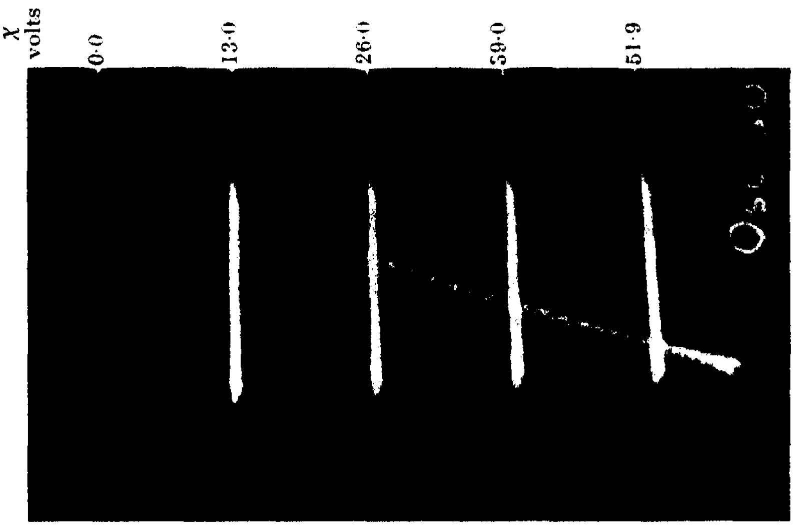

χ volts
0·0
13·0
26·0
39·0
51·9
Osc.

Fig. 4

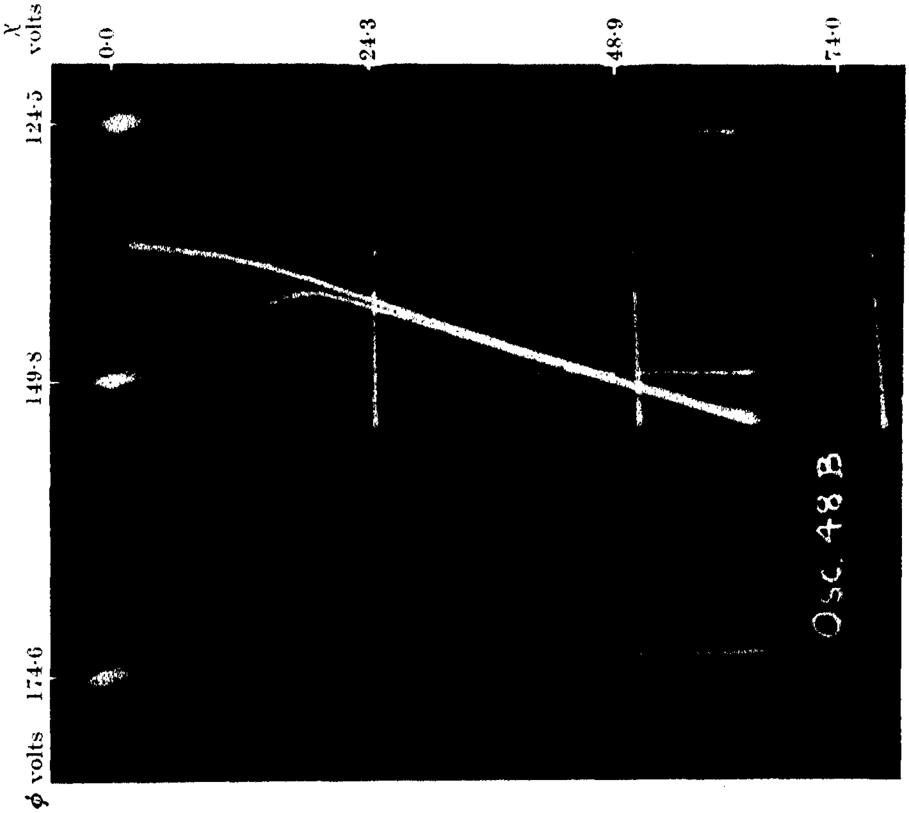

χ volts
0·0
24·3
48·9
74·0
124·5
149·8
174·6
ϕ volts
Osc. 48 B

Fig. 3

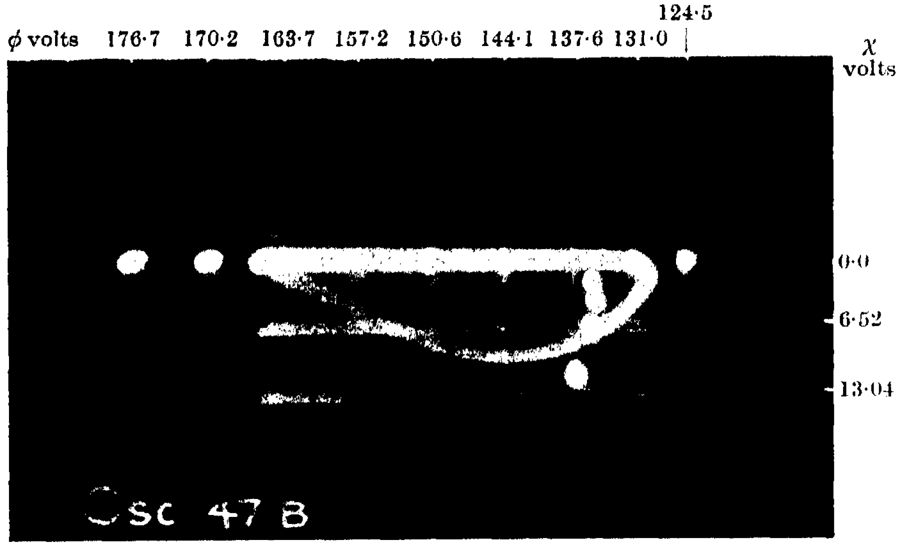

124·5
φ volts 176·7 170·2 163·7 157·2 150·6 144·1 137·6 131·0
χ
volts
0·0
6·52
13·04
Osc 47 B

Fig. 5

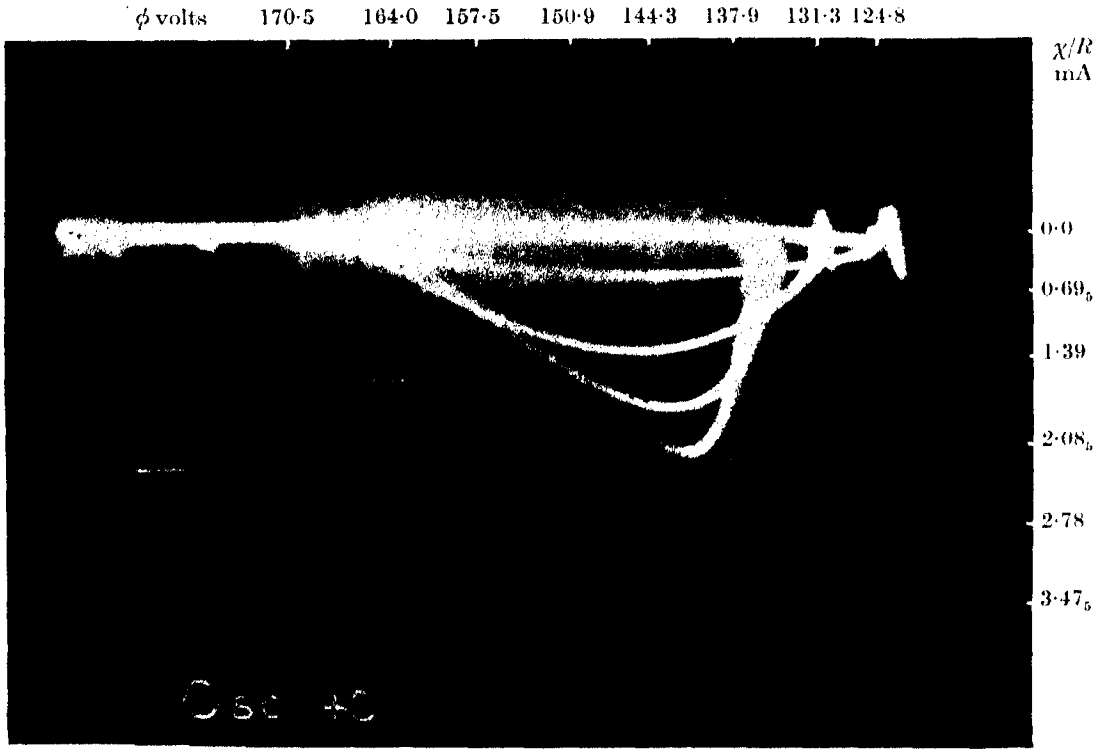

φ volts 170·5 164·0 157·5 150·9 144·3 137·9 131·3 124·8
χ/R
mA
0·0
0·69₅
1·39
2·08₅
2·78
3·47₅
Osc 4C

Fig. 6

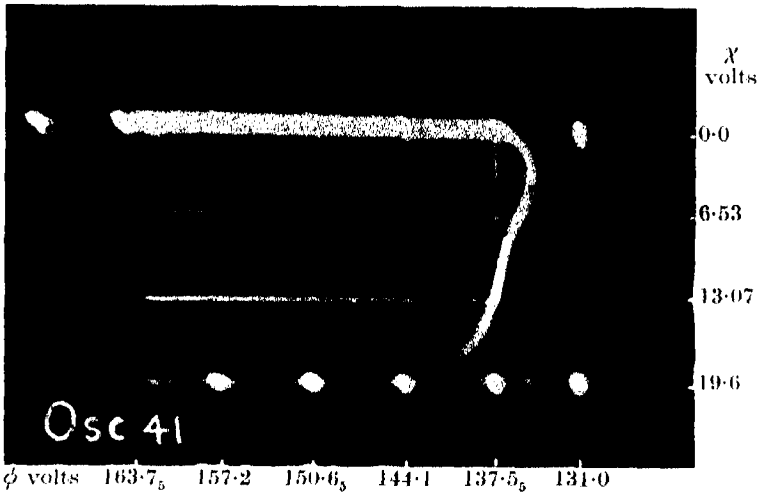
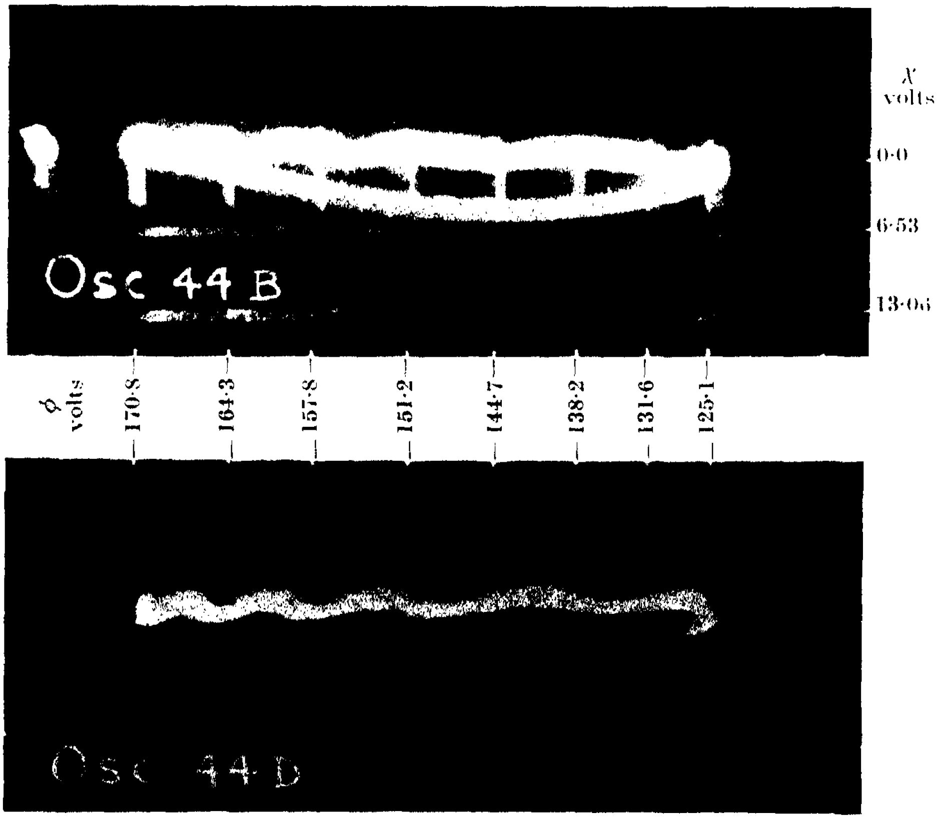


Fig. 7



Figs. 8 and 9

than its rise. At the higher frequencies the rates of rise and fall are more alike. It will be shown in the theory that this corresponds to a decrease in the importance of the "relaxational" terms.

TABLE III

Osc. No.	K 10^{-9} farad	Max. u mA	Min. ϕ V	n kc./sec.	nK 10^{-6} F/sec.	ν_0/n
43	60·3	4·2	—	Fickle	—	—
43	13·9	$2{\cdot}0_3$	—	0·131	1·82	$3{\cdot}9_6$
41	$13{\cdot}8_2$	1·84	134·7	$0{\cdot}130_8$	$1{\cdot}80_7$	4·00
43	$4{\cdot}8_0$	$1{\cdot}6_7$	—	$0{\cdot}35_7$	1·71	$2{\cdot}4_9$
43	$0{\cdot}9_8$	$1{\cdot}1_6$	131	$1{\cdot}21_9$*	1·59	—
43	$0{\cdot}0_7$	0·5	126	—	—	—
44	$0{\cdot}12_7$	$0{\cdot}5_3$	125	5·94	0·75	$0{\cdot}9_2$

* Or 4/3 of that; a stroboscopic whole-number uncertainty.

The last two columns of Table III are required for the theory given later.

At the higher frequencies the trace is more uniform in thickness, its shape is more like an ellipse and the equilibrium line lies nearer to the centroid of its area. See, for example, Osc. 44B and D (figs. 8 and 9, Plate 16). The ripple on the trace of Osc. 44 is a time mark of frequency 29·7 kc./sec. supplied from a triode oscillator. There are six time-mark waves in one cycle of the osglim. The trace frequency was therefore 5·94 kc./sec. It is seen that five-sixths of the osglim's period was spent near to zero current. In the most rapid part of the trace of Osc. 44 the stray current $E_1 d\chi/dt - F_1 d\phi/dt$ is of the order of 0·05 mA.

13·2—*Detailed Analysis of a Periodic Flash, namely, that of Osc.* 41—Ten points on the trace were located as being intersections with lattice lines, two others as being extremes and one was marked by an ink dot. These points were labelled α, β, γ, ..., ρ. See Table IV below.

The time t was computed from the current into the condenser and the voltage across it; a method previously used by Leyshon (1930). We have, in fact,

$$t = \int \frac{d\{C(\phi+\chi)+E_1\chi\}}{\dfrac{V-(\phi+\chi)}{S}-\dfrac{\chi}{R}}.$$

The stray capacitance E_1 was negligible, being swamped by C. This method of finding the time is satisfactory, except in the sudden breakdown, where $(\phi+\chi)$ is too nearly constant. But in this region between the points β and δ the trace was almost uniformly faint, with a slight minimum of strength near γ. Accordingly the time at the point γ was interpolated on the assumption that, between β and δ, time was proportional to the arc traversed by the light spot. The values of ϕ, χ, t are shown in Table IV. From these the current in the osglim was computed as

$$u = \frac{\chi}{R} + E_1 \frac{d\chi}{dt} - F_1 \frac{d\phi}{dt}.$$

The stray capacitances E_1, F_1 were assumed, in the light of later experience, to have been 10 and 30 $\mu\mu$F respectively; but, indeed, they are almost negligible.

TABLE IV

	α	β	γ	δ	ϵ	ζ	η
ϕ (volts)	$182{\cdot}8_6$	$157{\cdot}2_0$	$155{\cdot}2_8$	$150{\cdot}6_5$	$148{\cdot}0_8$	$144{\cdot}1_6$	$141{\cdot}0_1$
χ (volts)	$0{\cdot}0_0$	$5{\cdot}0_1$	$6{\cdot}5_3$	$10{\cdot}9_5$	$13{\cdot}0_8$	$15{\cdot}9_6$	$17{\cdot}2_4$
t (μsec.)	$< -10{\cdot}_8$	$0{\cdot}0$	$3{\cdot}7_5$	$13{\cdot}4$	$18{\cdot}5$	$28{\cdot}3$	$42{\cdot}5$
u (mA)	$0{\cdot}00_0$	$0{\cdot}54_3$	$0{\cdot}70_5$	$1{\cdot}17_5$	$1{\cdot}39_8$	$1{\cdot}70_4$	$1{\cdot}83_7$

	θ	κ	λ	μ	ν	ρ
ϕ (volts)	$137{\cdot}8_8$	$137{\cdot}5_5$	$136{\cdot}6$	$135{\cdot}7_4$	$134{\cdot}7_6$	$137{\cdot}5_5$
χ (volts)	$13{\cdot}0_6$	$12{\cdot}1_1$	$9{\cdot}9$	$6{\cdot}5_3$	$3{\cdot}5_3$	$0{\cdot}0_0$
t (μsec.)	$105{\cdot}_4$	$119{\cdot}_2$	$158{\cdot}_2$	$231{\cdot}_7$	$357{\cdot}_0$	[688]
u (mA)	$1{\cdot}39_0$	$1{\cdot}28_9$	$1{\cdot}05_3$	$0{\cdot}69_5$	$0{\cdot}37_6$	$0{\cdot}00_0$

Note to table—Although the absolute accuracy of ϕ is only 0·5 V, differences are significant to 0·05 V.

At each current u there are two portions of the oscillogram having opposite du/dt, say $\dot{u}_1$, $\dot{u}_2$, and different ϕ, say ϕ_1, ϕ_2. The bridge measurement of small oscillations suggested that $\dot{u}$ is proportional to the difference of potential from equilibrium ϕ_e for a given u. Accordingly ϕ_e was determined here from the assumption that

$$(\phi_1 - \phi_e)\,\dot{u}_2 = (\phi_2 - \phi_e)\,\dot{u}_1.$$

This is the only assumption borrowed from the results on small oscillations. The interpolations, which had to be made in order to obtain $\dot{u}_1$, ϕ_1, $\dot{u}_2$, ϕ_2, all at the same value of u, were effected graphically. The deduced values of ϕ_e are shown in Table V. If the reader will plot them on fig. 3 of Part I he will find that for the three smaller currents they agree closely with the static values taken from steady current dots on Osc. 66A. This agreement confirms the assumption. For the two larger currents of the present experiment ϕ_e progresses continuously, unlike the static values, which show a discontinuity where the glow jumps from one side of the kathode to cover the other side also. We may suppose that during the rapid changes of the present experiment the glow has not time to jump.

Each pair of observations yields also a value of $(\phi-\phi_e)/\dot{u}$ which is stated in henries in Table V. It is a slope-inductance and is in fair agreement with that shown in Part I, fig. 7 for small oscillations at a frequency of 5008 c./sec. In fact the present experiment (Osc. 41) gives

$$\frac{\phi-\phi_e}{du/dt} = 0{\cdot}017 + \frac{2{\cdot}7\times10^{-4}}{u},$$

TABLE V

	First member of pair		Common	Second member		Deduced quantities		Formula
Interval	du/dt amp. sec.$^{-1}$	ϕ V	u mA	du/dt amp. sec.$^{-1}$	ϕ V	ϕ_e V	$\frac{\phi-\phi_e}{du/dt}$ H	$0{\cdot}17 + \frac{2{\cdot}7\times10^{-4}}{u}$
β γ	$43{\cdot}_2$	$156{\cdot}2_4$	0·624	$-3{\cdot}_4$	135·5	$137{\cdot}0_1$	$0{\cdot}44_5$	0·450
γ δ	$48{\cdot}_7$	$152{\cdot}9_6$	0·940	$-5{\cdot}_1$	136·2	$137{\cdot}7_9$	$0{\cdot}31_2$	0·304
δ ϵ	$43{\cdot}_7$	$149{\cdot}3_6$	$1{\cdot}28_6$	$-6{\cdot}_7$	137·6	$139{\cdot}1_6$	$0{\cdot}23_3$	0·227
ϵ ζ	$31{\cdot}_3$	145·8	$1{\cdot}57_5$	−8·	138·7	$140{\cdot}1_3$	$0{\cdot}18_1$	0·188
Max. η	0	$141{\cdot}0_1$	$1{\cdot}83_7$	0	$141{\cdot}0_1$	$141{\cdot}0_1$	—	—

whereas for small oscillations the second constant was found to be $2{\cdot}8\times10^{-4}$ in close agreement; the unit being a volt-second. The first constant varies with frequency but is negligible at small currents.

The *important conclusion* is that the formula

$$\phi-\phi_e(u) = \left(a+\frac{b}{u}\right)\frac{dt}{du}$$

is a fair description not only of sinusoidal oscillations having an amplitude of 1 V, but also of very unsymmetrical oscillations during which $\phi-\phi_e$ exceeds 20 V.

The wave forms for current and voltage are shown in fig. 10. The periodic time, measured by a stroboscope, was $0{\cdot}0076_4$ sec. The rise of voltage during the absence of current has been calculated from $C\,d\phi/dt = (V-\phi)/S$ where V was the voltage of the driving battery. To fit these facts with the measured voltages at the points α and ρ, the time interval between them must be 688 μsec. This may be called, crudely, the duration of the flash. It is 0·090 of the periodic time.

An equilibrium exists. For if $d\phi/dt = 0 = du/dt$, we have simultaneously,

$$V = \phi+(S+R)\,u \quad \text{from Kirchhoff's laws,}$$

and

$$\phi = \phi_e(u) \quad \text{for the osglim.}$$

These equations were solved by the intersection of the straight line with fig. 5 of Part I. The solution is $\phi = 149\cdot_5$ V, $u = 0\cdot05_3$ mA. These values are shown by the thin lines crossing the wave form. On the oscillogram the equilibrium lies inside the loop. We may regard the motion of the light spot as a violent distortion of uniform motion in a circle; the equilibrium point is the point to which the centre has gone; the one wave form is a deformation of the sine curve, the other of the cosine curve.

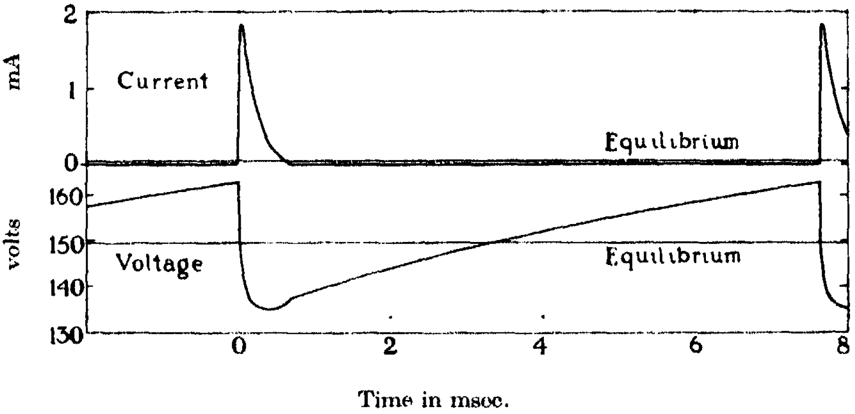


FIG. 10—Wave-forms of Osc. 41, see fig. 7, Plate 16.

13·3—*A Theory of Rapid Flashing*—It used to be supposed that a flash ended because the voltage had fallen to a "lower critical potential" (Taylor and Clarkson 1925). But at frequencies above 300 c./sec. that hypothesis is plainly untenable, and a new theory is required. The circuit is that shown in fig. 2. Let V be the voltage of the battery, K the capacitance, R and S resistances.

From Kirchhoff's laws for circuits it follows, whatever may be the properties of the osglim, that

$$\left(1+SK\frac{d}{dt}\right)\phi-V+(S+R)u+SKR\frac{du}{dt}=0. \tag{1}$$

The intrinsic equation has been shown to be approximately of the form,

$$\phi=\phi_c(u)+b\frac{d\log u}{dt}, \tag{2}$$

where

$$b=2\cdot7\times10^{-4}\ \text{V-sec.} \tag{3}$$

The elimination of ϕ between (1) and (2) yields,

$$SKb\frac{d^2\log u}{dt^2}+b\frac{d\log u}{dt}+SK\left\{R+\frac{d\phi_e}{du}\right\}\frac{du}{dt}+(S+R)\,u+\phi_e-V=0. \quad (4)$$

Equilibrium occurs at $u = u_0$, where

$$(S+R)\,u_0+\phi_e(u_0)-V=0. \quad (5)$$

It is assumed that $$V > \underset{u\to 0}{\text{Limit}}\,\phi_e(u). \quad (6)$$

The form of the $\phi_e(u)$ curve then shows that there is one and only one u_0.

It is not likely that any simple general solution of (4) can be found; but certain special cases are very instructive.

13·3·1—*Small oscillations*—Having found the value of u_0 from the curve, let $\phi_e(u)$ in that neighbourhood be represented approximately by its tangent line,

$$\phi = P+Qu, \quad (7)$$

where P and Q are functions of u_0, say $P(u_0)$, $Q(u_0)$.

Equation (4) may now be written

$$SKb\frac{d^2\log u}{dt^2}+\frac{d\log u}{dt}\{b+u\,.\,SK(R+Q)\}+(S+R+Q)\,u+P-V=0. \quad (8)$$

Let (8) be divided by $(P-V)$ and let the following dimensionless variables be introduced,

$$y=\left(\frac{S+R+Q}{V-P}\right)u, \quad (9)$$

$$x=t\sqrt{\frac{V-P}{SKb}}. \quad (10)$$

Then (8) simplifies to

$$\frac{d^2\log y}{dx^2}+\frac{d\log y}{dx}\{\alpha+\beta y\}+y-1=0, \quad (11)$$

in which α and β are dimensionless constants

$$\alpha=\sqrt{\frac{b}{SK(V-P)}},\quad \beta=\sqrt{\frac{SK(V-P)}{b}}\left(\frac{R+Q}{S+R+Q}\right). \quad (12)$$

The equilibrium value is $y = 1$; and, for small deviations from equilibrium,

$$y-1\simeq\log y = z \text{ say}. \quad (13)$$

So $$\frac{d^2z}{d^2x}+\frac{dz}{dx}(\alpha+\beta)+z=0. \quad (14)$$

An infinitesimal disturbance will be oscillatory if and only if

$$-2<(\alpha+\beta)<2. \tag{15}$$

It will grow if and only if $\alpha+\beta<0.$ (16)

This condition for self-starting may also be written

$$\frac{b}{SK(V-P)}<-\left(\frac{R+Q}{S+R+Q}\right), \tag{17}$$

which cannot be satisfied unless Q is negative.

If it were possible to have a small disturbance with α and β both near to zero, the period according to (14) would be $x = 2\pi$, that is

$$t = 2\pi\sqrt{\frac{SKb}{V-P}}. \tag{18}$$

13·3·2—*Maintenance of a Steady Range*—To discuss this we must return to the general equation (4). Let it be multiplied by $d\log u/dt$ and integrated with respect to t throughout a period T of an oscillation which is supposed to have attained steadiness. Four of the integrated terms vanish, leaving

$$0=\int_0^T\left(\frac{d\log u}{dt}\right)^2\left\{b+u\,.\,SK\left(R+\frac{d\phi_e(u)}{du}\right)\right\}dt. \tag{19}$$

Because b is positive, the integrand is positive when u is zero. The integrand is again positive when $u>0{\cdot}5$ mA, because $d\phi_e(u)/du$ is there positive. But there is an intermediate range of u, including part of the negatively sloping equilibrium line, for which the integrand may become negative, by suitable choice of S, K, R. Steadiness depends on the negative and positive portions together making the integral zero.

Van der Pol in 1926 suggested that the flashing of a neon lamp would probably be found to be an example of those "relaxation oscillations" which are described by his equation,

$$\ddot{v}-\epsilon(1-v^2)\dot{v}+v=0. \tag{20}$$

There is indeed a resemblance between (20) and (4) in so far as in both of them we have to do with a vibratory system in which the friction may be negative in the equilibrium position and become positive for large displacements of either sign. Apart from this important resemblance, the equation (4) is very different from (20), especially in the present of $\log u$. This effect will be considered in § 13·3·4.

13·3·3—*Discussion of the Oscillograms* 39–47—For these $V = 183{\cdot}_5$ V, $S = 6{\cdot}3_2 \times 10^6$ ohms, $R = 9397$ ohms, and K was adjustable. The equilibrium current, found from (5) with the aid of a graph of $\phi_e(u)$, was $u_0 = 5{\cdot}_4 \times 10^{-5}$ amp. See the lower curve of fig. 5 in Part I. The tangent line at this point has the coefficients $P = 158$ V, $Q = -1{\cdot}69 \times 10^5$ ohms; but with considerable uncertainty owing to the possible drift of equilibrium.

The condition (17) for self-starting becomes in this case $K > 0{\cdot}5 \times 10^{-10}$ F. This is in general agreement with the observations, for the least value of K at which a loop was photographed was $0{\cdot}_7 \times 10^{-10}$ F; and as K was decreased to this value the loop tended to flatten itself against the line $u = 0$, as though a limit were being approached (see Osc. 43, fig. 6, Plate 15).

None of these oscillations remained "small". For even in the flattest loop on Osc. 43 the ratio $u_{\text{max.}}/u_0 = 9 = y_{\text{max.}}$; so that u passed beyond the region in which $y - 1 \simeq \log y$ and $\phi_e(u) \simeq P + Qu$ are valid approximations. However, as a piece of empiricism, it is interesting to compare the observed frequency n with what it would have been if the oscillations had been (i) of that other relaxational type in which nK is constant, or (ii) of the small frictionless type of equation (18). The data for this comparison are set out in the last two columns of Table III. It is seen that as the terminal capacitance was approached, the oscillations became less large, nK did not remain constant, and the ratio of the observed period to $2\pi\sqrt{\{SKb/(V-P)\}}$ changed from 4·0 to 0·9. Concurrently the rates of rise and fall of current became more alike at a given value of the current. This last is a sign that as the terminal capacitance was approached the term in $d \log u/dt$ became less important.

13·3·4—*Large Oscillations with Slight Friction across a Straight* $\phi_e(u)$ *Graph*—In continuation of the policy of dividing the difficulties of solving equation (4) let us now neglect the terms which describe the supply of energy and which have already been studied in § 13·3·2. It is a commonplace remark among physicists that a moderate supply of energy to or from a vibrating system has very little effect on its period. Let us take this notion as a hint, but let us follow it sceptically; for it is derived from experience of vibratory systems unlike that now studied. Accordingly we simplify (4) and (11) to

$$\frac{d^2 \log y}{dx^2} + y - 1 = 0. \tag{21}$$

This equation is not merely an idealization. A disposition of the apparatus could exist for which the neglected terms would be really negligible. In the range $u < 3 \times 10^{-5}$ amp. the $\phi_e(u)$ graph is practically straight, so that formulae (11) and (12) apply, with the values

$$Q = -2{\cdot}8 \times 10^5 \,\text{ohm}, \quad P = 163{\cdot}5\,\text{V}.$$

See fig. 5 of Part I. There is room here for large oscillations, because $\log u$ is unbounded at $u = 0$. By a suitable choice of R, S, V, K it is possible to make both α and β near to zero, and yet to preserve the self-starting condition $\alpha + \beta < 0$. For example, let $R = 2{\cdot}7 \times 10^5$ ohm, so that $|Q/R| \doteqdot 1$. Let $S = 10^6$ ohm, $V = 168{\cdot}5$ V, $K = 10^{-8}$ F, then formulae (12) give $\alpha = 0{\cdot}07$, $\beta = -0{\cdot}14$. Under circumstances of this sort we may expect that equation (21) would give a good approximation to the wave-form and period. But this setting of the apparatus differs considerably from that used for Osc. 41–47.

Equation (21) has been made the subject of a separate study by Carson and Richardson (unpublished). It is there shown that, as the range of oscillation increases, the crests of the waves become narrow and steep and they become separated by wide troughs in which y approaches extraordinarily near to zero. See the values of $y_{\text{max.}}$ and $y_{\text{min.}}$ in Table VI. That is to say, the crests will resemble isolated flashes; yet their isolation will not be complete, for the current will not entirely cease.

TABLE VI—PROPERTIES OF THE INTEGRALS OF

$$d^2 \log y/dx^2 + y - 1 = 0$$

$y_{\text{max.}}$	$\to 1$	7	25	35	100
$y_{\text{min.}}$	$\to 1$	0·0064	$3{\cdot}47 \times 10^{-10}$	$2{\cdot}2 \times 10^{-14}$	$3{\cdot}8 \times 10^{-42}$
Period	$\to 2\pi$	$8{\cdot}4_5$	$14{\cdot}_6$	$17{\cdot}_1$	$28{\cdot}_5$
$\dfrac{\text{Period}}{2\pi}$	$\to 1$	$1{\cdot}3_4$	$2{\cdot}3_3$	$2{\cdot}7_2$	$4{\cdot}5_4$

In the experiments the current certainly remained at less than 5×10^{-5} amp. during the greater part of the flashing cycle. But there is evidence that at frequencies above 1000 c./sec. the current never quite ceased. For in other experiments, in which the current certainly passed through zero because it changed sign, the oscillograms showed more or less flickering. A similar unsteadiness was observed for the largest and slowest loop in Osc. 43; attempts to measure the frequency with a stroboscope failed for that reason. But above 1000 c./sec. the oscillograms were perfectly steady, as would be expected if the current had not to start from zero.

We are now able to compare the observed periods with what they would have been if (21) were an accurate substitute for (4). That is to say, we can now make allowance for the large range of oscillation by using a factor

T/T_0 computed from equation (21); T_0 being the period for small oscillation, T that for an oscillation having its $y_{\max.}$ equal to the observed $u_{\max.}/u_0$, where u_0 is the equilibrium current of $5\cdot_4 \times 10^{-5}$ amp. The theoretical period, when thus corrected, is

$$\frac{T}{T_0} 2\pi \sqrt{\frac{SKb}{V-P}} = \tau, \text{ say.}$$

The required data are extracted from Table III and from the computed integrals of (21) (Carson and Richardson unpublished) and are set out in Table VII. By comparison with Table III it is seen that the correcting factor T/T_0 brings the ratio of observed to theoretic periods much nearer to unity for the larger amplitudes. The remaining deviations from unity may be attributed to the neglect of the relaxational terms, of the curvature of the $\phi_e(u)$ graph, and of small terms in the intrinsic equation of the osglim.

TABLE VII

Observed frequency kc./sec.	$\dfrac{u_{\max.}}{u_0}$	$\dfrac{T}{T_0}$	$\dfrac{\text{Observed period}}{\tau}$
$0\cdot131$	$37\cdot4$	$2\cdot8_1$	$1\cdot4_2$
$0\cdot130_8$	$34\cdot1$	$2\cdot6_8$	$1\cdot4_9$
$0\cdot357$	$30\cdot_9$	$2\cdot5_7$	$0\cdot9_7$
$1\cdot21_9$	$21\cdot_5$	$2\cdot1_4$	$0\cdot7_5$
$5\cdot94$	$9\cdot_8$	$1\cdot4_9$	$0\cdot6_2$

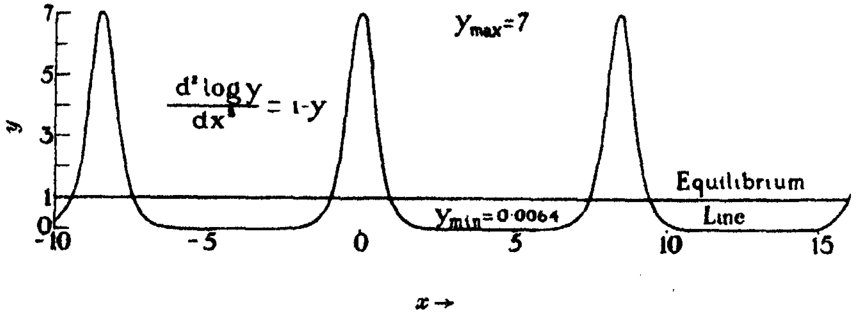


FIG. 11.

For giving a clear view of the peculiarities of the wave-form defined by equation (22) a less violent departure from the cosine-curve is more pleasant. So for illustration $y_{\max.} = 7$ has been chosen, and is shown in fig. 11. It is roughly comparable with the wave-form deducible from Osc. 44. The broad flat valleys and sharp isolated peaks are reminiscent of a number of physiological waves (Starling 1926) and of some in psychology (Flugel 1928; Richardson 1929). The equal slope on the two sides of the crests in fig. 4 *is due to the neglect of the relaxational terms in equation* (4).

13·4—Conclusion on Relaxational Oscillations—The equation

$$\phi - \phi_e(u) = b\frac{d \log u}{dt},$$

which was derived from small oscillations, is upheld as a good approximate description under the very different circumstances of flashing. The constant b is about $2{\cdot}7 \times 10^{-4}$ V-sec. The range of frequency was from 0·131 to 6 kc./sec. The range of current was from 2 mA downwards.

The author gratefully acknowledges grants for apparatus made to him by the Government Grant Committee in the years 1929 and 1930.

14—SUMMARY

Oscillations of various sizes and of various frequencies mostly below 1000 c./sec. were recorded by a cathode ray oscillograph with full attention to the correction of instrumental errors. One coordinate gave voltage, the other gave current and the time was obtained either from time marks superposed on the trace or from the charge of and current into a condenser. The results were compared with the empirical intrinsic equation

$$\phi - \phi_e(u) = \left(a + \frac{b}{u}\right)\frac{du}{dt},$$

which had been induced in Part I as a description of small oscillations, ϕ being voltage, u current, a and b constants and $\phi_e(u)$ the equilibrium voltage. It was found that this equation remained a true description in oscillations that were very unsymmetrical and not small. The constant a was negligible at small currents; this was fortunate, because its value decreased with the frequency. The value of b was $2{\cdot}_7 \times 10^{-4}$ V-sec.

At this stage the type of reasoning changes from inductive to deductive. The empirical intrinsic equation is taken as a starting point and the behaviour of the osglim, when connected with battery, condenser and resistance in the well-known flashing circuit, is deduced mathematically. If the static

characteristic $\phi_e(u)$ is approximated by its tangent line, the current in the osglim is proportional to y where

$$\frac{d^2 \log y}{dx^2} + \frac{d \log y}{dx}\{\alpha + \beta y\} + y - 1 = 0,$$

in which x is proportional to the time, and α and β are constants. If $-2 < \alpha + \beta < 0$ and initially $y = 1$, an oscillation begins to grow. It belongs to the general class of relaxational oscillations in the wider sense. But, owing to the presence of the logarithm, the wave-form differs remarkably from that of the relaxational oscillation first analysed by B. van der Pol. Brief disturbances are separated by long pauses, during which the current becomes too small to be measured, but never quite ceases. Physiological and psychological analogues are suggested.

It will be shown in a later communication that some small terms in the intrinsic equation, which are not necessary for the explanation of flashing, and which for simplicity have been omitted from the present theory, do under other circumstances become important.

REFERENCES

Anson, H. St. G. and Pearson, S. O. 1922 *Proc. Phys. Soc.* **34**, 175–6 and 204–12.
Carson, J. and Richardson, L. F. Unpublished.
Flugel, J. C. 1928 "Practice, Fatigue and Oscillation." *Brit. J. Psychol. Monog. Suppl.* **4**, No. 13.
Leyshon, W. A. 1930 *Proc. Phys. Soc.* **42**, 157–69.
Pearson, S. O. and Anson, H. St. G. 1922 *Proc. Phys. Soc.* **34**, 175–6, 204–14.
Pol, B. van der 1926 *Phil. Mag.* **2**, 978–92.
Richardson, L. F. 1929 *J. Gen. Psychol.* **2**, 339, Fig. 6.
— 1933 *Proc. Phys. Soc.* **45**, 135–41.
— 1935 *Proc. Phys. Soc.* **47**, 258–62.
— 1937 *Proc. Roy. Soc.* A, **162**, 293–316.
Starling, E. H. 1926 "Principles of Human Physiology", figs. 424*b*, 516. London: J. and A. Churchill.
Taylor, J. and Clarkson, W. 1925 *Proc. Phys. Soc.* **37**, 130–41.

Stark Effect in Iron and the Contrast with Pole Effect

By S. F. Panter, Ph.D., *Demonstrator in Physics*
and J. S. Foster, F.R.S., *Macdonald Professor of Physics, McGill University, Montreal*

(*Received* 31 *March* 1937)

[Plates 17, 18]

Introduction

In the operating conditions of the iron arc there are a number of influences which may measurably affect the universally accepted standards of wave-lengths. Since this phenomenon appears to have some complexity, it is important to learn what part may be played by pure Stark effect.

Already it is known that an increase of pressure within the source causes small shifts in the wave-lengths of most iron lines (pressure shift), which according to Babcock (1928) is always toward the red. This effect appears to be different from the so-called pole effect which consists of a relative change of wave-length at the pole as compared with that at the centre of the arc. For it is found that in the latter effect the shift is toward the red for some lines and toward the violet for others. In recent years the pole effect has been rather generally accepted as a pure Stark effect due to electric fields developed by the relatively large drop of potential near the pole. This view has indeed been taken as a result of earlier examinations of the Stark effect for iron in which the displacements of a few lines were reported to have the same sign as their pole effects.

In the present more extended investigation of the Stark effect at much greater field strength only partial agreement with the earlier work has been reached. It will be shown that the sign of the Stark shifts for a few lines are opposite to that of their pole effects. As regards the magnitudes of the two effects there is apparently no relation. Indeed the pole effects are of the same order of magnitude for lines which (1) show no Stark effect and (2) show maximum effect. On the basis of new observations on a few hundred lines one is forced to conclude that the pole effect is certainly not identical with the pure Stark effect, and indeed that between these two there may be little connexion.

The known fact that the pressure shift is not the same for all iron lines led Gale and Adams (1912) to classify the iron spectrum in four pressure groups *a*, *b*, *c*, *d* in order of increasing magnitudes of the shifts. An additional group *e* was later found by St. John and Ware (1912). Recent precise measurements on 130 lines (6896–6678 A) by Babcock (1928), who used an international arc with vacuum arc comparison, have shown that the pressure shift is always toward the red, even for the lines of group *e* for which the pole effect is toward the violet.

The pole effect was discovered by Goos (1913), who regarded it as a kind of pressure effect. This view has been considered disproved by the work of Babcock (1928). Later work by several investigators has failed to reveal its origin with certainty.

The Stark effect in iron has received very little attention. A few lines in the green region were examined in a field of 29,800 V/cm. by Takamine (1919). While all displacements were small compared with the effects in most other elements investigated, it appeared that nine lines were shifted to the violet and two to the red. Takamine pointed out that the direction of the shift for each line agreed with that of the pole effect. More recently Adam (1932) has made a further study of the relation between pole effect and Stark effect, and for this purpose has examined the Stark effect in the region 4000–5000 A at a maximum field of 28,500 V/cm. No photographs were published, but it is stated that the lines 4250, 4260, 4872, 4891, 4920, 4957 were shifted toward the red, the displacements of the last three lines being of the order of 0·5 A. Adam concluded that there was complete agreement as regards the sign of the displacements in the pole and Stark effects.

In the present research 225 lines in the region 3400–6500 A have been photographed in electric fields of more than 150,000 V/cm. As already stated, the results show only partial agreement with the above.

CLASSIFICATION OF THE ARC SPECTRUM OF IRON

The normal arrangement of the energy levels of an atom is always of first importance in studies of the Stark effect, since displacements depend not only upon the type of level immediately concerned but especially upon its separation from neighbouring levels, and the character of the latter. Classifications within the normal iron spectrum are due mainly to Laporte; more than 2400 lines have now been accounted for as combinations between established energy levels. Many lines not yet detected in the laboratory have been predicted by the application of selection rules to known iron terms and have been found among previously unidentified solar lines.

Extensive analyses of the arc spectrum of iron have been published by Burns and Walters (1929) and by Catalan. From the terms compiled by

Bacher and Goudsmit (1932), the energy-level diagram (fig. 1) has been constructed, with notation as modified by Moore (1932). The prefixes *a*, *b*, *c*, etc., are used to denote successive low even terms of any one *L*-value and multiplicity, written in the order of increasing energy; *e*, *f*, *g*, ... represent the similar high terms. The odd terms are denoted successively by the prefixes *z*, *y*, *x*, etc., *z* being the lowest. A number of terms have not yet been classified; they are designated by numbers, and the usual symbol (°) is added to those of the odd set. The classification of this spectrum is by no means complete.

EXPERIMENTAL

The aim in designing a Lo Surdo discharge tube for investigation of the Stark effect in spectra of metals of relatively high melting points is to obtain a source of light in which the action of the discharge produces metallic vapour in the neighbourhood of the cathode. In the early experiments a high current (up to 40 mA) appeared to be essential to the fulfilment of the above requirements. The relatively low electric fields which were developed were due to the use of a cathode of large diameter.

The inner diameter of the tube surrounding the cathode in the present investigation is only 1·3 mm., the current only 2 mA, while the maximum electric field attained is in excess of 150,000 V/cm.

It has already been observed by Anderson (1917) that the spectra of metals used as the cathode are excited to greater brightness when the residual gas is either air or oxygen than when it is hydrogen. In the present investigation air proved to be preferable to either oxygen, nitrogen or helium. The pressure of air used was from 2·0 to 2·5 mm. of mercury.

The complete discharge tube is shown diagrammatically in fig. 2. To avoid loss of heat from the cathode a 3 l. pyrex bulb was used for the main body of the discharge tube. This served also as a reservoir, preventing large variations in pressure due to oxidation of the cathode during operation. The iron cathode *C*, of about 1·3 mm. in diameter, fits into the quartz capillary tube *D*. The upper end of the cathode is turned down as shown in the diagram in order not to touch the walls of the capillary tube. This isolation of the cathode makes it possible to put a high voltage across the terminals of the tube without causing instability of the discharge. The quartz tube *G*, which is terminated by the capillary *D*, is fitted in the glass tube *I* by grinding, and *I* is sealed to the lower part of the ground-glass joint *J*. The glass tube *F*, which is also sealed to the ground-glass joint, is terminated by tungsten wire sealed at each end. The purpose of *F* is to provide an electrical contact for

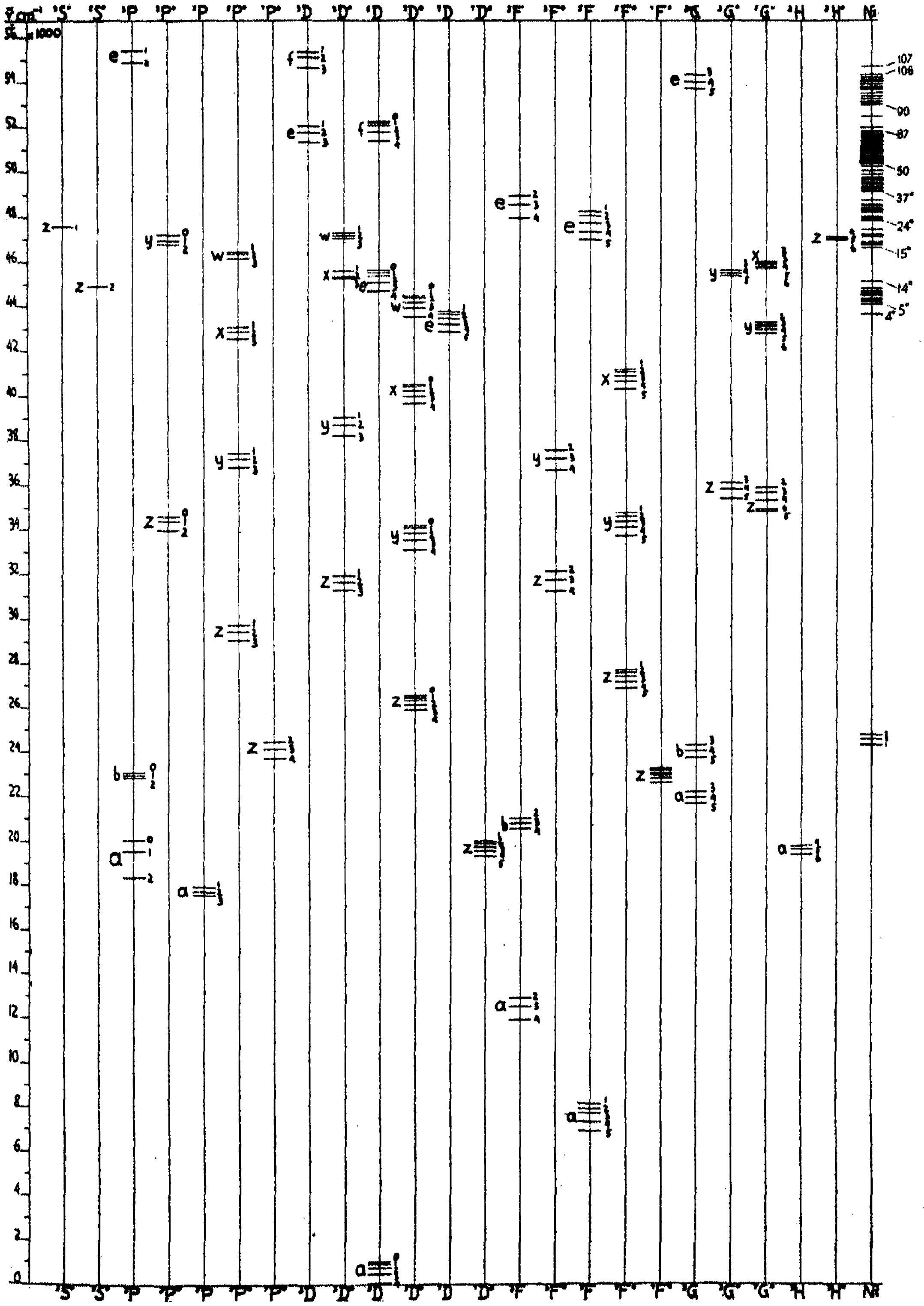


FIG. 1—Energy level diagram of the iron atom.

the cathode through the wire *H* and spring *E*. Within the tube *F* there is atmospheric pressure to prevent any discharge from taking place along the wire *H*. A narrow slit of 2 mm. length was cut along the axis of the capillary tube. The bright light from the beam traversing Crookes's dark space leaves the discharge tube through the quartz window *B*. The aluminium anode *A*

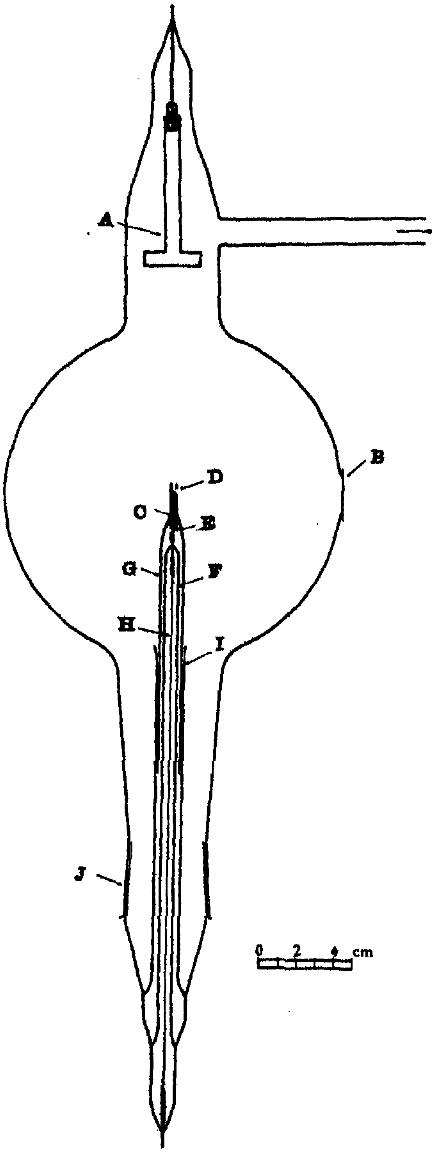
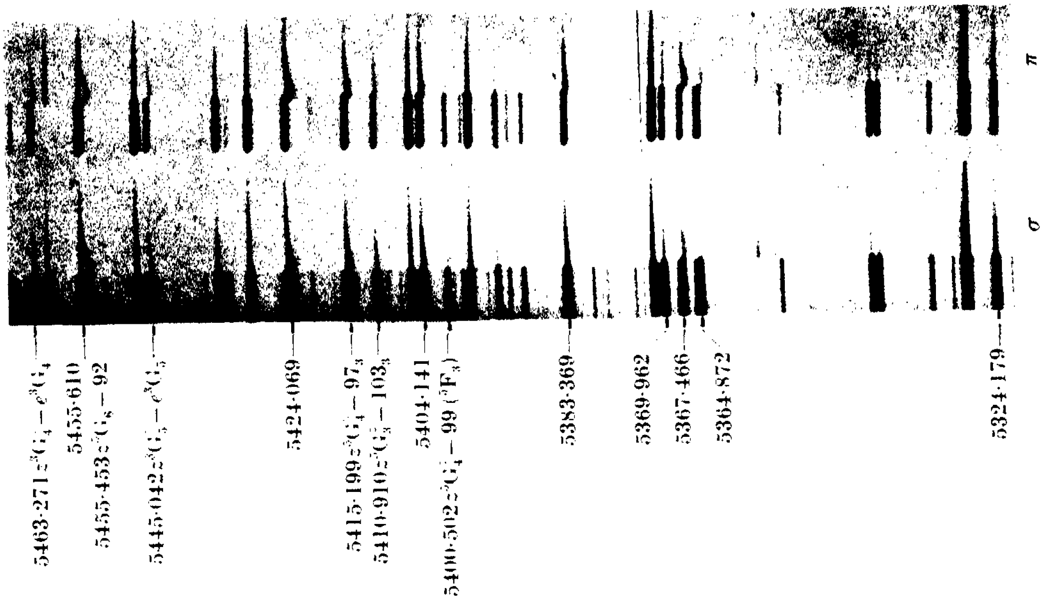
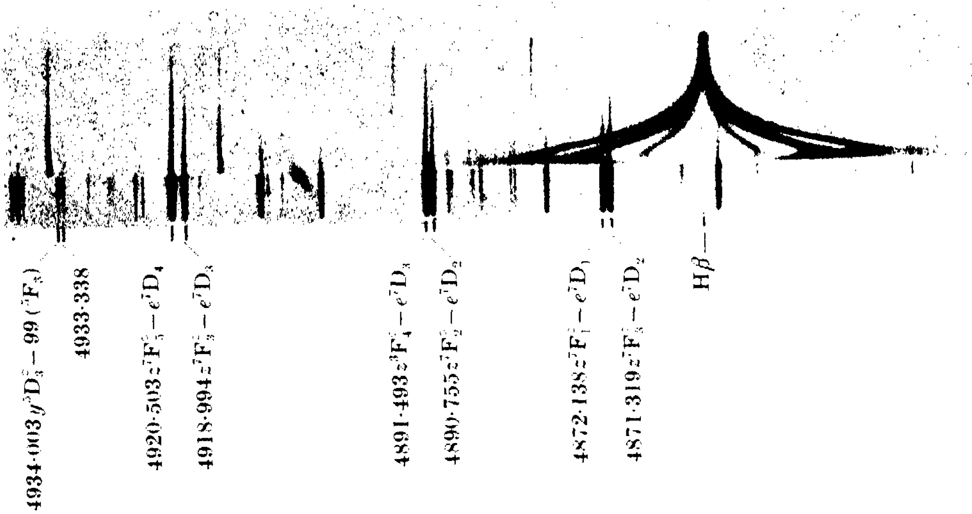


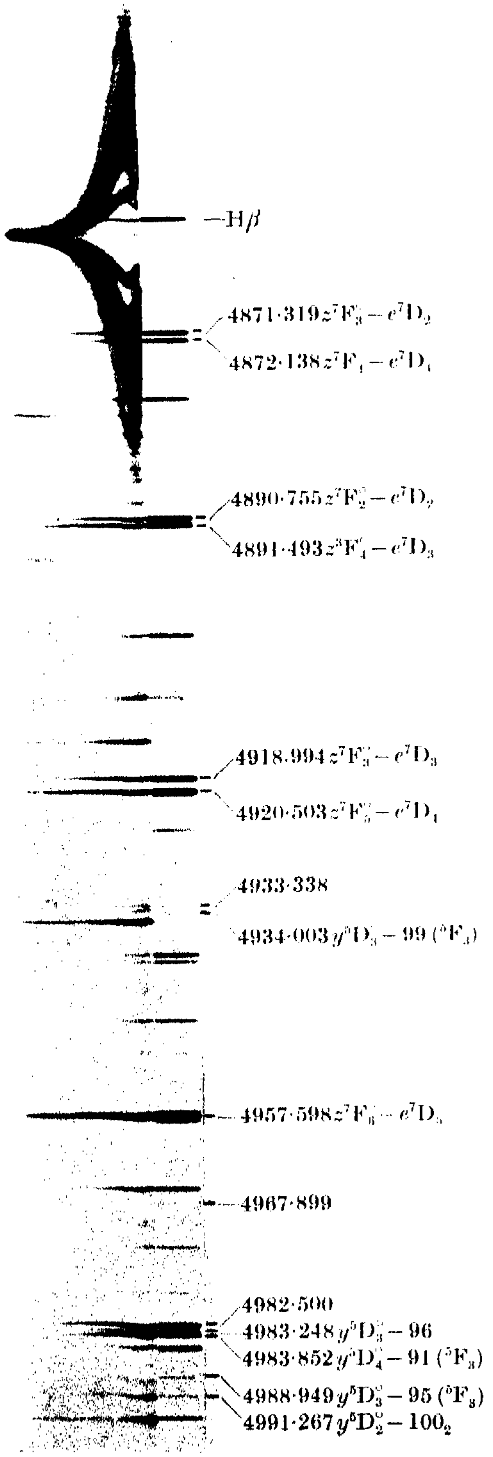
Fig. 2—Discharge tube.

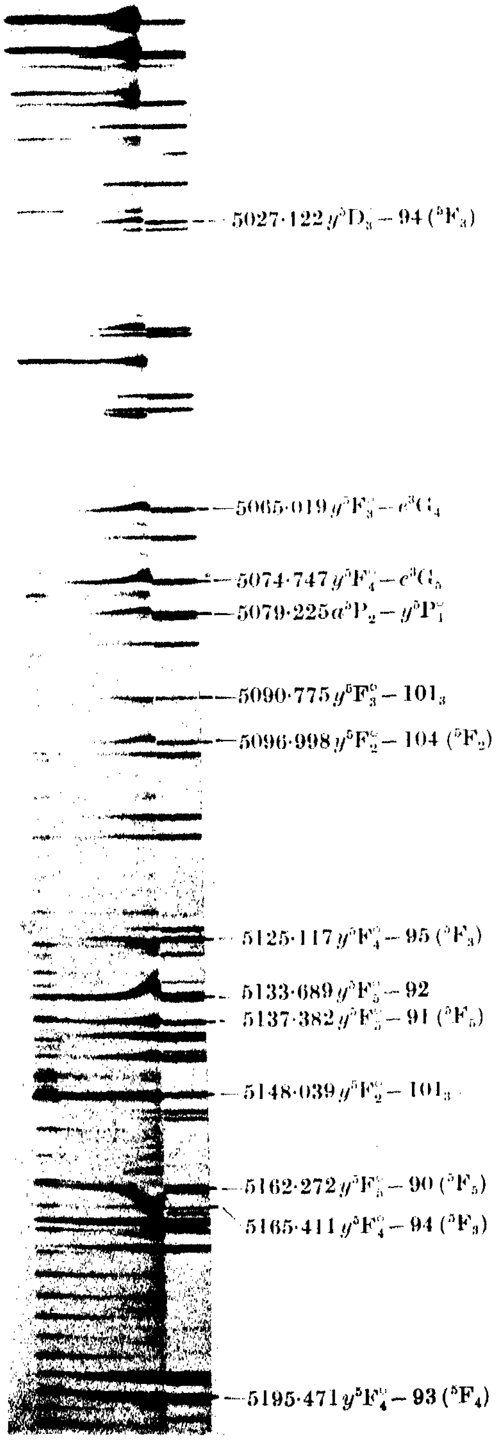


a. Maximum electric field, 115,000 V/cm.



b. Maximum electric field, 115,000 V/cm.


—Hβ
4871·319 $z^7F^\circ_3 - e^7D_2$
4872·138 $z^7F_1 - e^7D_1$
4890·755 $z^7F^\circ_2 - e^7D_2$
4891·493 $z^3F^\circ_4 - e^7D_3$
4918.994 $z^7F^\circ_3 - e^7D_3$
4920·503 $z^7F^\circ_5 - e^7D_4$
4933·338
4934·003 $y^5D_3 - 99\ (^5F_4)$
4957·598 $z^7F_6 - e^7D_5$
4967·899
4982·500
4983·248 $y^5D^\circ_3 - 96$
4983·852 $y^5D^\circ_4 - 91\ (^5F_3)$
4988·949 $y^5D^\circ_3 - 95\ (^5F_3)$
4991·267 $y^5D^\circ_2 - 100_2$


5027·122 $y^5D^\circ_3 - 94\ (^5F_3)$
5065·019 $y^5F^\circ_3 - e^3G_4$
5074·747 $y^5F^\circ_4 - e^3G_5$
5079·225 $a^5P_2 - y^5P^\circ_1$
5090·775 $y^5F^\circ_3 - 101_3$
5096·998 $y^5F^\circ_2 - 104\ (^5F_2)$
5125·117 $y^5F^\circ_4 - 95\ (^5F_3)$
5133·689 $y^5F^\circ_5 - 92$
5137·382 $y^5F^\circ_5 - 91\ (^5F_5)$
5148·039 $y^5F^\circ_2 - 101_3$
5162·272 $y^5F^\circ_5 - 90\ (^5F_5)$
5165·411 $y^5F^\circ_4 - 94\ (^5F_3)$
5195·471 $y^5F^\circ_4 - 93\ (^5F_4)$

may be unscrewed when necessary, and removed from the discharge tube through the ground-glass joint *J* by means of a pair of tongs.

The discharge tube was mounted on a wooden frame in front of the spectrograph and attached to the vacuum system by means of a ground-glass joint. The vacuum system was of the simplest type, consisting only of a McLeod gauge and a reservoir for air. It was connected to a Hyvac pump used for evacuating the system.

The D.C. high-potential source used in this experiment was a split transformer kenotron rectifier, capable of delivering a maximum current of 100 mA at a potential of 10,000 V. The potential across the tube was measured with an electrostatic voltmeter.

For the visible region a large six-prism glass spectrograph was used, the dispersion of which varied from about 8 A/mm. at 5500 A to 2·5 A/mm. at H_δ. The construction of this spectrograph has been described elsewhere (Foster 1929). For the ultra-violet region between 4000 and 3300 A, a quartz spectrograph was employed which has been recently designed by Foster (1936). The dispersion varies from 6 A/mm. at 4000 A to 4·5 A/mm. at 3500 A.

The light from the source passed through a quartz Wollaston prism and was focused on the slit of the spectrograph by means of a condensing lens, arranged to give a magnification of about 1·5. The light polarized with its electric vector parallel to the field was then passed through a quartz quarter-wave plate, which rotated the plane of polarization of the light through 90°, and hence the loss of light by reflexion was greatly reduced.

The plates used were Eastman spectroscopic types I–F, I–D, I–J, each in its most sensitive region and Eastman "40" in the ultra-violet.

Considering the fact that the discharge tube is rather large and difficult to adjust during operation, it was thought advisable to make all the adjustments of focusing before the beginning of the run. An ordinary electric bulb was placed behind the discharge tube, and the image of the slit in the quartz capillary tube was brought into focus on the slit of the spectrograph by the condensing lens, the proper position of which was previously determined. Thus the illuminated slit of the quartz capillary served as a source of light and made it unnecessary to make any adjustments of focusing of the tube during the course of an actual run.

As soon as the pressure was reduced to about 2 mm. of mercury, the high potential was applied and the tube was ready for a run. The voltage across the terminals was about 10,000 V, and the current 2 mA. The discharge was always steady and of bluish colour. Exposure times varied from 1 to 3 hr., depending upon the pressure used and potential applied. Generally speaking this type of discharge tube proved to be stable and satisfactory so far as high fields and intensity of the light source are concerned.

For the purpose of classification and measurement of displacements of the affected lines, a comparison iron arc spectrum was put at the end of each run. A mirror was placed behind the condensing lens and the image of the iron arc was focused on the slit of the spectrograph without disturbing the optical system. A mask on the slit served to limit the light of the iron arc to the proper positions on the plate.

To prepare the tube for another run, the quartz tube was removed and boiled in HCl to dissolve iron oxide and iron deposited on the walls of the capillary tube. The same capillary may be used for several runs when necessary. It may easily be replaced by a new piece sealed on the main quartz tube *G* by means of a very hot oxygen flame.

EXPERIMENTAL RESULTS

Thirty-three lines were found to be displaced in fields of 150,000 V/cm. or less; of these eight were displaced toward the red and the rest toward the violet. The maximum field in each exposure was determined from the separation of strong components of the Balmer lines H_α, H_β or H_γ which always appeared on the plates owing to water vapour in the tube. In photographs of the ultra-violet region the field was obtained from the displacements of two nitrogen lines 4100, 4110 A which appear in the overlapping regions of the two spectrographs.

In Table I the first three columns give the wave-length, wave number and term combination of the lines; they are taken from a paper of Burns and Walters (1929). The next two columns give the pressure classification and pole effect taken from the results of St. John and Babcock, which are included for the sake of later comparisons. Columns six and seven give the displacements of the parallel and perpendicular components of the lines in cm.$^{-1}$. For many lines only the sign of the small displacement is given.

Plate 17 (*a*) is an enlargement of a plate taken with a double-image prism and showing displacements at a maximum field of 115,000 V/cm. in the longer wave-length region 5465–5325 A. The two sections of Plate 18 are enlargements from a single plate taken at a maximum field of 150,000 V/cm. without a double-image prism, and covering the region 5200 A to H_β. Although very good plates were obtained in this region with a double-image prism, the above plate is reproduced owing to the appearance of certain faint lines not recorded elsewhere. Plate 17 (*b*) shows the H_β region in a maximum field of 210,000 V/cm. in which additional lines are displaced.

Table I—Displacements of Iron Arc Lines between H_{α} and H_{β}

λ in A	$\bar{v}$ in cm.$^{-1}$	Term combination	Pressure group	Pole effect in A	Stark effect in cm.$^{-1}$		Field in V/cm.
					p	*e*	
5463·271	18298·981	$z^3G^{\circ}_4-e^3G_4$	*e*	−0·016	−1·74	−1·13	115,000
5455·453	325·204	$z^5G^{\circ}_6-92$	*e*	−0·016	−3·95	−3·78	115,000
5445·042	360·244	$z^3G^{\circ}_5-e^3G_5$	*e*	−0·020	−1·76	−1·60	115,000
5424·069	431·235	—	*e*	−0·027	−2·24	−1·68	115,000
5415·199	461·423	$z^5G^{\circ}_4-97_3$	*e*	−0·030	−2·04	−0·88	115,000
5410·910	476·057	$z^5G^{\circ}_3-103_3$	*e*	−0·026	−1·19	−0·65	115,000
5404·141	499·199	—	*e*	−0·025	−1·27	−0·68	115,000
5400·502	511·664	$z^5G^{\circ}_4-99$ $(^5F_3)$	*e*	−0·011	−0·89	−0·89	115,000
5383·369	570·579	—	*e*	−0·025	−1·131	−0·55	115,000
5369·962	616·944	—	*e*	−0·020	−small	0	115,000
5367·466	625·598	—	*e*	−0·025	−2·52	−1·55	115,000
5364·872	634·605	—	*e*	−0·028	−1·62	−(diffuse)	115,000
*5195·471	242·190	$y^5F^{\circ}_4-93$ $(^5F_4)$	*e*	—	—	−small	150,000
5165·411	19354·168	$y^5F^{\circ}_4-94$ $(^5F_3)$	*d*	+0·010	−small	−small	115,000
5162·272	365·936	$y^5F^{\circ}_5-90$ $(^5F_5)$	*d*	+0·030	+5·76	+4·02	81,000
*5148·039	419·477	$y^5F^{\circ}_2-101_3$	?	—	—	+small	150,000
5137·382	459·760	$y^5F^{\circ}_5-91$ $(^5F_5)$	*d*	+0·014	−small	−small	112,000
5133·689	19473·761	$y^5F^{\circ}_5-92$	*e*	−0·045	−3·94	−3·65	112,000
5125·117	506·330	$y^5F^{\circ}_4-95$ $(^5F_3)$	*d*	+0·020	+0·98	+0·72	81,000
5096·998	613·943	$y^5F^{\circ}_2-104$ $(^5F_2)$	*e*	−0·016	−1·31	−1·61	112,000
5090·775	637·918	$y^5F^{\circ}_3-101_3$	*d*	+0·010	+small	0	112,000
5079·225	682·572	$a^5P_2-y^5P^{\circ}_1$	*e*	−0·013	−1·08	−1·43	81,000
5074·747	699·939	$y^5F^{\circ}_4-e^3G_5$	*e*	−0·022	−1·52	−1·70	112,000
5065·019	737·776	$y^5F^{\circ}_3-e^3G_4$	*e*	−0·029	−1·40	−1·16	81,000
5027·122	886·569	$y^5D^{\circ}_3-94$ $(^5F_3)$	*d*	+0·014	−0·98	−small	112,000
4991·267	20029·423	$y^5D^{\circ}_2-100_2$	*d*	+0·030	+small	+small	81,000
4988·949	038·726	$y^5D^{\circ}_3-95$ $(^5F_3)$	*d*	+0·018	+small	+small	81,000
4983·852	059·223	$y^5D^{\circ}_4-91$ $(^5F_5)$	*d*	+0·020	−small	−small	112,000
4983·248	061·651	$y^5D^{\circ}_3-96_2$	*d*	+0·020	+small	+small	112,000
4982·500	064·664	—	*d*	+0·026	+small	+small	112,000
*4967·899	123·634	—	*e*	−0·016	—	−small	150,000
*4934·003	261·880	$y^5D^{\circ}_3-99$ $(^5F_3)$	?	—	—	−small	150,000
*4933·338	20264·612	—	*e*	−0·006	—	−small	150,000

* These lines were photographed without a double-image prism.

Discussion

The first point to be noted is that the pole effect is not identical with the Stark effect in the arc spectrum of iron. Owing to the following new evidence, the earlier assumption that the two effects are the same is not upheld. (1) The four *d*-lines λ 5165·411, 5137·382, 5027·122 and 4983·852 are found with definite Stark displacements toward the violet, whereas the observed pole shifts for all of these lines are toward the red. (2) There is, in fact, no clear relation between the magnitudes of displacements of individual lines in Stark effect and pole effect respectively. For example, the pole effects for the lines λ 5463 and 5455 are the same, while the Stark effect observed for the second line is double that of the first. (3) A great many lines which are undisplaced in Stark effect have about the same pole effects as other lines which are appreciably affected by electric fields. While the above evidence is clear, some possibilities in Stark effects demand closer consideration. With reference to the first point, it is possible and indeed not uncommon to find a line which shifts to one side of its normal position in low electric fields ("pole effect"?) and to the other side in high fields ("Stark effect"). Moreover, with special reference to the second point, lines in a complex spectrum are usually not all displaced by electric fields according to the same law, and hence one should not expect the ratio of the displacement to be necessarily independent of the field strength. It might be thought that these observations invalidate points (1) and (2) in the above argument; but apparently they do not apply to the iron spectrum. Here the Stark displacements display an uncommon regularity, being roughly proportional to the square of the field and devoid of any signs of the irregularities just cited. We therefore conclude that the pole effect is not identical with the Stark effect.

As a second point, it may be mentioned that the origins of Stark effects in the iron spectrum are much less evident than in most spectra. The energy-level diagram (fig. 3) shows that the levels which are appreciably affected by the field are relatively high and close together. It should be noted that the scale used in the representation of displacements is much larger than that used to indicate separations of normal levels. The apparent crossing of levels in the diagram arises from this change of scale and is not actually realized. According to present classifications, all the terms corresponding to the displaced levels are *even*. This feature is worthy of note, since it means that there can be no interactions between the affected terms themselves to give the Stark effects. Acceptance of the present consistent interpretation of the normal spectrum apparently forces the prediction of odd levels interspersed with the even levels which are perturbed. In other words, one must postulate odd levels (e.g. E_z, ψ_z) to produce non-vanishing elements in the matrix z given by

$$\int \psi_a z \psi_z^* \, dv,$$

where ψ_a is the eigenfunction for an even state. Since it is well known that the odd and even levels would "repel" each other in the field, the observed Stark displacement of known levels (fig. 3) may be interpreted to mean that one or more odd levels are present at 53,060, 53,650 and 54,000 cm.$^{-1}$

approximately. Since the sum of the Stark displacements of the odd and even levels must be zero, the observed displacements of known terms rather suggest that the odd levels will not be greatly displaced by the field; but small shifts may be considered probable. Owing to the absence of even terms in the middle section of the energy level diagram fig. 1 (25,000–42,000 cm.$^{-1}$), one cannot expect that evidence for the predicted terms will be found in the region covered by the present investigation; and indeed no such evidence appears. There are a great many even terms below 25,000 cm.$^{-1}$, however, and it therefore seems probable that some unclassified lines to the violet of 3400 A might show significant Stark effects and have the right wave-lengths to serve as evidence for the odd terms in question.

Table II has been constructed from the 225 photographed iron lines with the aid of the present classification given by Burns and Walters (1929). Of the ninety-eight initial levels associated with these lines, sixteen are affected in electric fields and the rest are not displaced in fields up to 150,000 V./cm. It appears, however, from Plate 18 (*b*) that the group of levels e^7D is slightly displaced in fields of over 200,000 V/cm.

From the known behaviour of the ninety-eight initial levels in electric fields (Table II) one may predict Stark effects for a good many lines not yet investigated. Such predictions are based on the reasonable assumption that no effect exists in the final states and have proved to be quite accurate during the course of the present work.

It has already been noted that Takamine (1919) reported the displacements of eleven iron lines in the green region in a maximum field of 29,800 V/cm. According to his observations the line λ 5324·179 is displaced toward the red, but reference to Plate 17 (*a*) will show without any doubt that this line is undisplaced even in a field of 115,000 V/cm. The remaining ten lines are found on the present plates with displacements in the direction reported by Takamine. In some cases, however, the shifts are detected only in fields of the order of 100,000 V/cm.

Displacements of the order of 0·5 A at 28,500 V/cm. reported by Adam (1932) for the lines λλ 4957, 4920, 4891 are not confirmed. Under the present dispersion a change in wave-length of the above magnitude is easily detected; but it is not found for these lines at fields of 150,000 V/cm. as may be learned from Plate 18. The lines are only slightly affected at 210,000 V/cm. (Plate 17 (*b*)). Of the remaining displaced lines reported by Adam, λ 4872 shows no shift except at 210,000 V/cm.; λλ 4260, 4250 are not displaced at 135,000 V/cm.

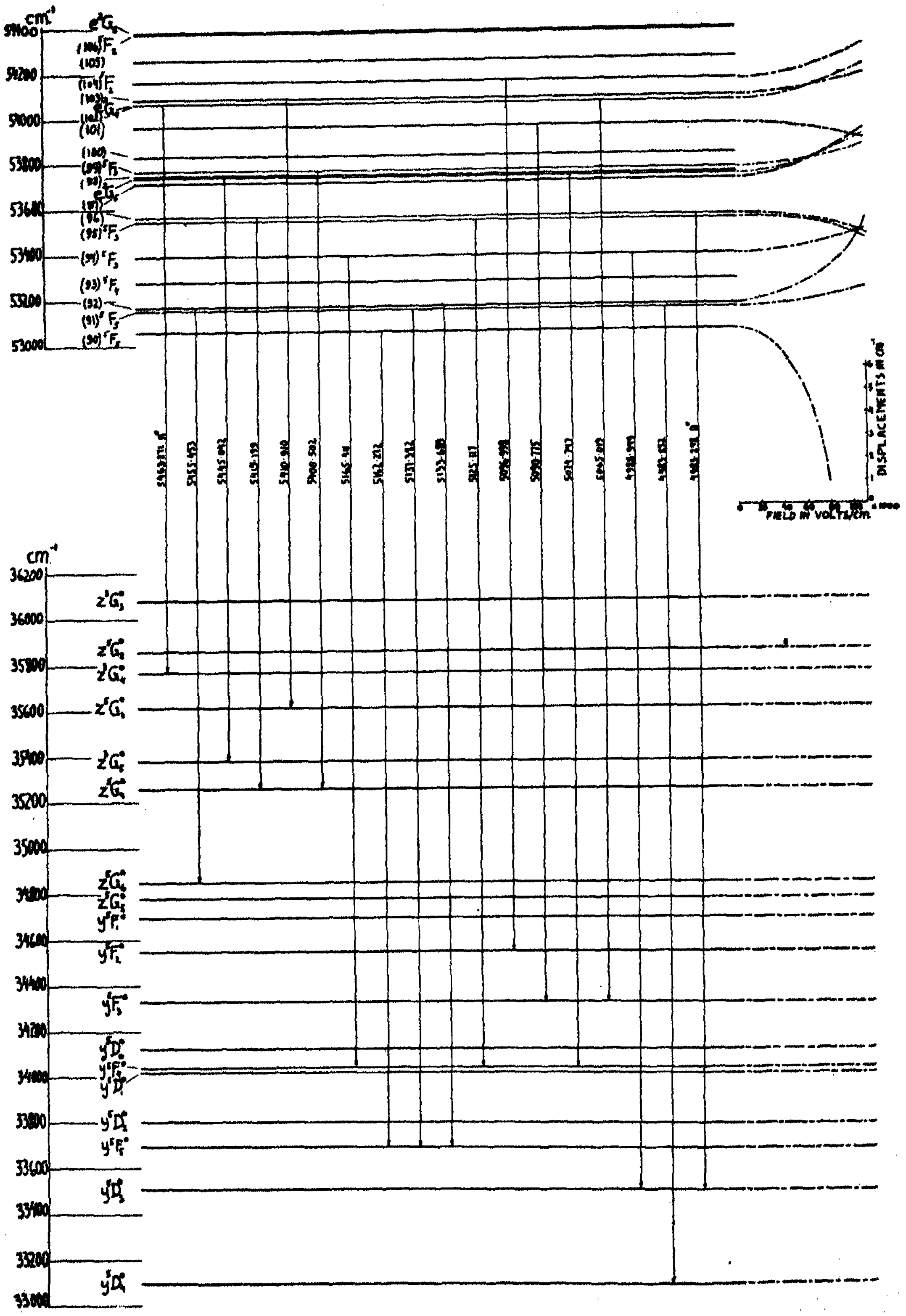


FIG. 3—Energy level diagram showing levels displaced by electric fields.

TABLE II—DISPLACED AND UNDISPLACED LEVELS

Symbol	Displaced J	Undisplaced J	? J
z^5S°		2	
z^3P°		2	
			1
			0
z^5P°		3	
		2	
		1	
y^5P°		3	
		2	
	1		
e^3D		3	
		2	
		1	
f^3D			3
			2
		1	
z^3D°		3	
		2	
		1	
y^3D°		3	
		2	
		1	
x^3D°		3	
			2
			1
w^3D°		3	
		2	
			1
e^5D		4	
		3	
		2	
		1	
		0	
f^5D		4	
			3
		2	
			1
			0
z^5D°		4	
		3	
		2	
		1	
		0	
y^5D°		4	
		3	
		2	
		1	
		0	
x^5D°		4	
		3	
			2
			1
			0
e^7D		5	
		4	
		3	
		2	
		1	
z^7D°			1
			2
			3
		4	
			5
z^3F°		4	
		3	
		2	
y^3F°		4	
		3	
		2	
e^5F		5	
		4	
		3	
			2
			1
z^5F°		5	
		4	
		3	
		2	
		1	
y^5F°		5	
		4	
		3	
		2	
		1	
e^3G	5		
	4		
			3
z^3G°		5	
		4	
		3	
z^5G°		6	
		5	
		4	
		3	
		2	
z^3H°		6	
			5
			4

Unassigned levels

Displaced	Undisplaced
104	78
103	70
101	65
100	61
99	60
97	47
96	45

Unassigned levels

Displaced	Undisplaced
95	42
94	36
93	—
92	—
91	—
90	—

SUMMARY

A modified Lo Surdo discharge tube has been designed suitable for investigations of Stark effects in the spectra of metals of high melting points.

The Stark effect in the arc spectrum of iron has been observed in the region 6500–3400 A at a number of field strengths ranging from 100,000 to 200,000 V/cm.

Two hundred and twenty-five lines were photographed, thirty-three of which were found to be displaced.

The Stark effect consists of simple small displacements of the lines. Some polarized components were observed.

The relation between the pole effect and the Stark effect is discussed, and the conclusion drawn that the pole effect cannot be identical with the Stark effect as has been assumed by many writers.

Stark displacements of lines not yet examined are predicted from the known effects on ninety-eight initial terms.

A general theoretical explanation of the observations requires the introduction of hitherto unknown odd terms of specified values.

REFERENCES

Adam, M. 1932 *Ann. Phys., Lpz.*, **15**, 568–611.
Anderson, J. A. 1917 *Astrophys. J.* **46**, 104–16.
Babcock, H. D. 1928 *Astrophys. J.* **67**, 167–88.
Bacher, R. F. and Goudsmit, S. 1932 "Atomic Energy Levels." New York: McGraw-Hill Book Co.
Burns, K. and Walters, F. M. 1929 *Pub. Allegheny Obs.* **6**, 159–211.
Foster, J. S. 1929 *Proc. Roy. Soc.* A, **122**, 599–603.
— 1936 *Canad. J. Res.* A, **14**, 173–6.
Gale, H. G. and Adams, W. S. 1912 *Astrophys. J.* **35**, 10–47.
Goos, F. 1913 *Z. wiss. Phot.* **12**, 259.
Moore, C. E. 1932 *Astrophys. J.* **75**, 222–73.
St. John, C. E. and Ware, L. W. 1912 *Astrophys. J.* **36**, 14.
Takamine, T. 1919 *Astrophys. J.* **50**, 23–41.

The Stark Effect in Hydrogen and Deuterium

By J. Stuart Foster, D.Sc., F.R.S., *Macdonald Professor of Physics*
and Hawley Snell, Ph.D., *Research Associate, McGill University*

(*Received* 31 *March* 1937)

[Plates 19, 20]

Introduction

Since hydrogen and deuterium present observable differences in the fine structure of the Balmer series, it was thought advisable to examine the Stark effects with special reference to intensities.

The photographs have been taken from a Lo Surdo source containing a mixture of hydrogen and deuterium. Plate 19 shows the results of the analyses of $H_\alpha D_\alpha$, $H_\beta D_\beta$, $H_\gamma D_\gamma$ and $H_\delta D_\delta$. It may be noted that the intensity pattern for D_α is asymmetric and that it is nearly the mirror image of the asymmetric pattern for H_α. Although gradually less prominent in later members of the series, this feature is not entirely absent from $H_\gamma D_\gamma$ as published in *Nature* (Foster and Snell 1934). Moreover, the contrasting asymmetries of D_α, H_α persist to some extent in analyses (Plate 20 (1)) from a canal ray source specially constructed with opposing beams, and operated at a pressure enormously in excess of usual values.

There are other intensity relations of some interest. The hydrogen line in the pair $H_\gamma D_\gamma$ has less relative strength than it has in $H_\beta D_\beta$ taken from the same plate (Plate 20 (2)). In other words, under the conditions of the experiment, the intensities in the Balmer series of deuterium do not fall off so rapidly as they do in hydrogen. Confining the attention now to a single pair of lines, e.g. $H_\gamma D_\gamma$, the deuterium intensities are relatively greater in electric fields than at zero field (Plate 20 (2)). While this point is beyond doubt, one must not include in the evidence the central component of D_γ upon which a molecular line falls. It is clear that some of the intensity relations are a result of collisions of the second kind, and it appears probable that others originate in self absorption.

The more accurate measurements of Stark displacements for H_γ and H_β (Foster 1924, 1925) and for H_α (McRae 1931) have shown variations from the regular Epstein spacing which are outside experimental error. McRae showed that the irregularities in H_α are nearly those to be expected if the conditions between fine structure and Stark effect are as predicted (Schlapp

1928), and the fine-structure separations are carried over into high fields. On our new plates the effect of fine-structure separations is seen as the main cause of the irregular spacing of Stark components in the four Balmer lines of deuterium, and the earlier displacements in hydrogen are confirmed.

While all hydrogen components are diffuse, it is established in the present investigation that certain selected components are rendered particularly diffuse, and usually in an asymmetric manner.

EXPERIMENTAL INVESTIGATION

Lo Surdo Method—The greatest part of the experimental study was made with the Lo Surdo source developed in this laboratory. Before admission of the hydrogen gases, the tube was thoroughly roasted and run for some time with helium gas in contact with charcoal cooled by liquid air. The greater part of the helium was then pumped out and replaced by hydrogen and deuterium in approximately equal proportions. During the final runs, the combined partial pressures of the hydrogen isotopes was certainly greatly in excess of that of helium. The total pressure varied with the field strength required for a good analysis of the line under investigation. It was 0·9 mm. for $H_\alpha D_\alpha$ and $H_\beta D_\beta$, and 1·6 mm. for $H_\gamma D_\gamma$ and $H_\delta D_\delta$. The corresponding voltages on the tube varied from 14,000 V for $H_\alpha D_\alpha$ to 9000 V for $H_\delta D_\delta$. No difficulty was encountered in exciting the Balmer lines of deuterium or in maintaining a good intensity throughout the exposure.

A brief consideration shows that a clear Stark analysis of a given Balmer pair may be obtained at a certain minimum external field. As is well known, the quantum number n serves as a splitting factor in *hydrogen* Stark effect. Each Stark pattern of a Balmer line ($n \rightarrow 2$) is therefore made up of groups of three components (end-groups are incomplete) with separations $(n-2)$ Epstein units. The separation of the groups may be equal to this, or greater, but not less. Resolution will be satisfactory when the deuterium components lie midway between hydrogen components of each group, and will not be much improved by the use of higher fields. On this basis, the minimum field for satisfactory analysis is the one for which

$$(\tfrac{1}{2}n-1) \text{ Epstein units} = \text{separation of Balmer pair.}$$

The numerical values are: for $H_\alpha D_\alpha$, 131,000 V/cm.; $H_\beta D_\beta$, 87,000 V/cm.; $H_\gamma D_\gamma$, 65,000 V/cm.; $H_\delta D_\delta$, 52,000 V/cm. Our experimental values, given on Plate 19, differ a little from the above.

The spectrograph is a large glass instrument (Foster 1930) in which six prisms were required for the analyses of $H_\alpha D_\alpha$ and $H_\beta D_\beta$, while three were sufficient for $H_\gamma D_\gamma$ and two for $H_\delta D_\delta$. The dispersion varied from 12·2 A/mm. for H_α to 7·0 A/mm. for H_δ. Exposures lasted from 30 to 60 min.

The intensities of components may be estimated from the photographs reproduced in Plate 19. Under our experimental arrangement, we considered it improbable that the physical conditions could be repeated with the accuracy necessary to duplicate exactly any set of intensities, and hence have been content with qualitative results which show, under identical conditions, sufficiently interesting contrasts between the two isotopes.

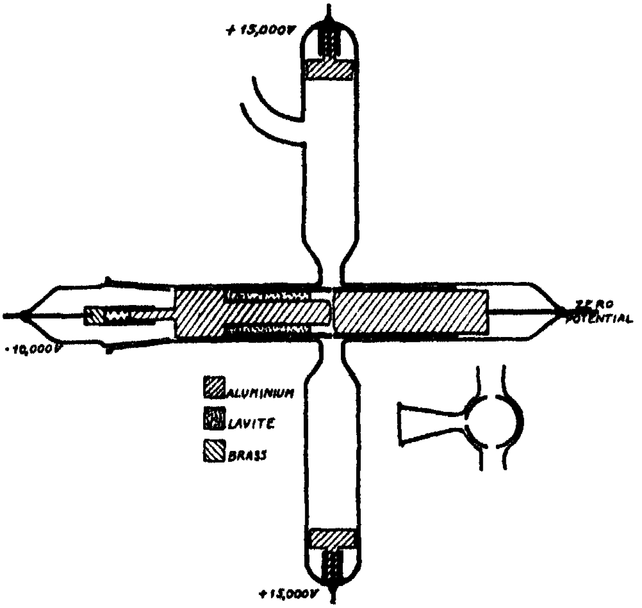


FIG. 1

Canal-ray Investigation with Cross-Tube—A few photographs of $H_\alpha D_\alpha$ have been taken to see if the asymmetric Stark patterns for the isotopes would appear in a canal-ray source at high pressure. For this purpose, we have employed a source the essential features of which were developed with Dr. D. R. McRae for another research. Two directly opposing canal-ray beams meet at a common cathode, where an additional electrode is introduced to enable one to apply an electric field perpendicular to the beams.

The present form of cross-tube is shown in cross-section in fig. 1. The two anodes are connected to the positive terminal of a Kipp "proton" apparatus. The common cathode is an aluminium tube with wall 1½ mm. thick. Rectangular slits ½ × 2 mm. are punched at opposite ends of a diameter to allow the canal-ray beams to pass through the cathode. Each beam may be seen to project a distance of about 2 cm. on the other side. The external field is applied between two field plates, one of which is an aluminium cylinder accurately fitted into the tubular cathode, and held by a set screw. This unit, combining two cathodes and one field plate, is grounded. The second field plate is held rigidly in the cathode tube by an accurately turned lavite sleeve. The arrangement is such that the field plates are kept parallel—an important feature in the investigation. Although space between the plates is 1 mm. or less, the plates are placed carefully outside the canal beams. A ground-glass joint facilitates the removal of the cathode and plates for cleaning and resurfacing. A diameter of 1 cm. for the cathode leads to convenient dimensions for the assembly. The light passes through a third slit in the cathode placed at 90° from the canal-ray slits as illustrated separately in fig. 1. It continues through a side window, condensing lens, and Wollaston prism to the spectograph.

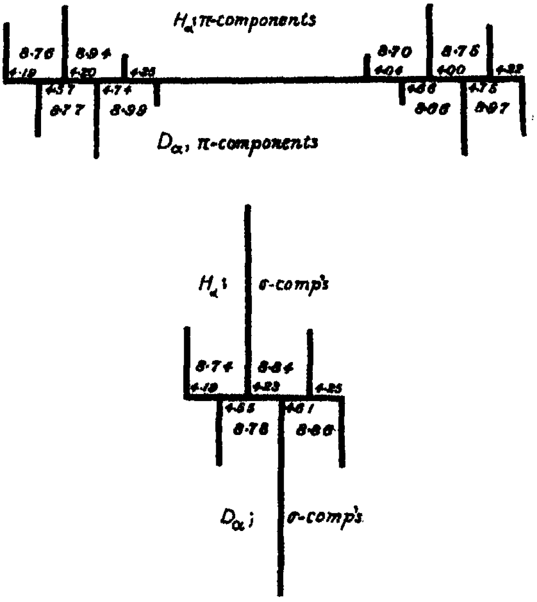


FIG. 2

The cross-tube will operate at pressure as high as 0·85 mm., and give light of intensity comparable to that of a Lo Surdo source. It is believed that the pressure and light intensity may be increased beyond the present limits. The field potential remains very steady provided care has been taken to prevent the canal beams from striking the field plates.

Irregularities in Displacements—In figs. 2 and 3 we have summarized the mean results of twenty independent measurements of each interval between the components of the $H_\alpha D_\alpha$ and $H_\beta D_\beta$ photographs reproduced in Plate 19. In these diagrams the wave-length increases to the left, and it will be noticed that the figures immediately above the horizontal line represent the H-D separations. All figures are in cm.$^{-1}$.

The irregularities in the intervals which are to be expected on the basis of fine structure are listed in Table I. Here $(4-3)\,v$, (r) represents the interval between components 4 and 3 on the violet (red) side, and the figures represent the differences between intervals as deduced from figs. 2 and 3.

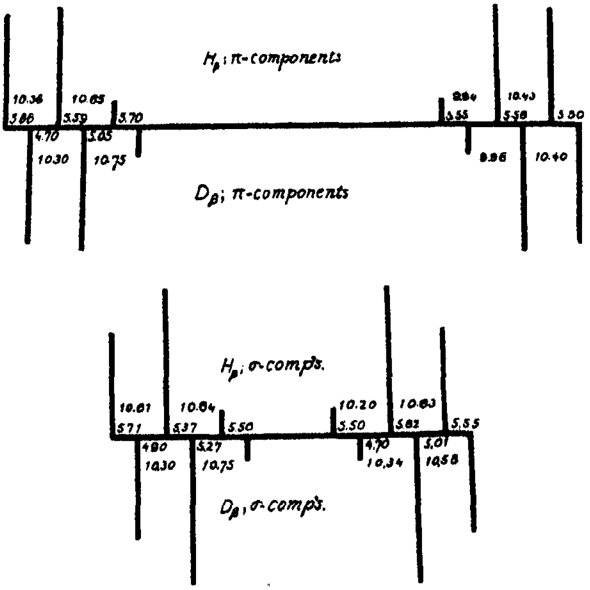


FIG. 3

It will be noticed that practically all the differences are positive in sign. This means that in both H and D the small and large intervals occur where they are expected from fine structure considerations. The numerical agreement may be impaired by the fact that only one plate was measured for each line, and (especially in $H_\beta D_\beta$) certain components are very diffuse. Taken as a whole, however, the evidence is strongly in favour of attributing the irregularities in the displacements to the effects of fine structure. This conclusion is supported by measurements of five plates of $H_\gamma D_\gamma$. In these plates the following irregularities are found in both H_γ and D_γ:

π: $(5-2)\,r > (5-2)\,v$.

σ: $(13-10)\,v > (13-10)\,r$; $(10-3)\,r > (10-3)\,v$; $(3-0)\,v > (3-0)\,r$.

These again are in agreement with the expected effect of fine structure. In addition we should have $(18-15)\,v > (18-15)\,r$. Here our measurements are contradictory but inconclusive, owing to the greater differences of these wide components and some confusion with molecular lines.

TABLE I

	Difference between intervals			
Intervals compared	Theory	H (F. & S.)	D (F. & S.)	H (McRae)
$H_\alpha D_\alpha$, π: $(3-2)\,r-(3-2)\,v$	0·21 cm.$^{-1}$	0·24 cm.$^{-1}$	0·33 cm.$^{-1}$	0·27 cm.$^{-1}$
$(4-3)\,v-(4-3)\,r$	0·16	−0·01	0·20	0·10
$(4-3)\,v-(3-2)\,v$	0·19	0·05	0·31	0·15
$(3-2)\,r-(4-3)\,r$	0·19	0·18	0·22	0·20
σ: $(1-0)\,v-(1-0)\,r$	0·19	0·10	0·08	—
$H_\beta D_\beta$, π: $(10-8)\,v-(10-8)\,r$	0·19	0·07	0·10	—
$(8-6)\,r-(8-6)\,v$	0·19	0·71*	0·79*	—
$(8-6)\,r-(10-8)\,r$	0·19	0·29	0·45	—
$(10-8)\,v-(8-6)\,v$	0·19	0·49	0·44	—
σ: $(6-4)\,v-(6-4)\,r$	0·19	0·02	0·28	—
$(4-2)\,r-(4-2)\,v$	0·19	0·44	0·41	—
$(4-2)\,r-(6-4)\,r$	0·19	0·03	0·45	—
$(6-4)\,v-(4-2)\,v$	0·19	0·43	0·24	—

* This large difference is not supported by measurements of other plates.

Relatively Diffuse Components—Small variations of the external field during the exposure account for a general blurring of components which increases with the displacement. Clearly superposed on this, however, is a selective diffuseness not previously reported. It is illustrated best in H_β, Plate 19, where the components $-6(\sigma)$; $6(\pi)$ are decidedly broad as compared with their mates. Since the component $6(\pi)$ and the relatively sharp component $4(\sigma)$ are associated with the same initial level, the difference in their profiles cannot arise from effects in the initial state. On the other hand, the diffuse components just mentioned are associated with the same final state. A more detailed examination of all Balmer patterns reveals the fact that almost all relatively diffuse components are produced in transitions to the lowest final state.

This effect is not due to a general variation of the external field. Neither is it due to an incorrect adjustment of the spectrograph or to the fine structure of Stark components (Schlapp 1928). The observations allow one to consider a surprisingly selective effect on the final state concerned.

Beginning with the discovery of the effect and continuing up to the present time, much material has been amassed to prove that the observed hydrogen Stark intensities may vary from theory by amounts many times greater than the experimental error. Moreover, the character of the variation changes greatly with the physical conditions of the discharge. In the canal-ray source it depends, for example, upon the relative orientation of the beam and the applied field—whether parallel, anti-parallel, or perpendicular—as well as upon the voltage, pressure, nature of canal rays and of residual gas.

The apparent departures from theory are not, however, limited to canal-ray sources. In this laboratory, McRae (1931) examined a Lo Surdo source containing pure hydrogen gas, and observed a type of asymmetric pattern for H_α which could not be explained by any loading of the initial states. The asymmetry was less at the lower pressures.

More recently a rather smooth transition from the hydrogen Stark intensity patterns normally characteristic of the canal ray source to those characteristic of the Lo Surdo method (Thornton 1935) has been obtained. This has been realized by the simple process of increasing the gas pressure in a suitably designed canal-ray tube to values far above normal.

On the other hand, there are published quantitative measurements of average relative Stark intensities in hydrogen (Foster and Chalk 1929) which agree rather well with the calculated (Schrödinger 1926) relative probabilities of transition. The experiments were carried out with a Lo Surdo source containing a small amount of hydrogen mixed with helium or neon at a total pressure of one to three millimetres. Under these conditions an excited hydrogen atom may exist and radiate with comparatively little disturbance from its immediate neighbours, and the radiation is not appreciably absorbed on its way to the spectrograph. Although no other series of measurements has shown such good agreement with theory, the correct point of view seems to be that no other experimental arrangement has so nearly realized the conditions assumed in the theoretical treatment.

We now find that the asymmetric Stark-Lo Surdo intensity pattern characteristic of H_α from *pure hydrogen* (McRae 1931) is retained in photographs from mixtures of the isotopes, while its mirror image represents rather well our finding for D_α. Most physical conditions are here constant, yet the asymmetries are quite different. While searching for an explanation of this stimulating fact—without denying a theory which had received adequate confirmation in the same laboratory—it was found that selective absorption within the source (Foster 1934) apparently serves as the main satisfactory supplement to earlier ideas. The actual effects are not simple, however, as is seen from the fact that collisions of the second kind must be invoked to explain the relatively weak outer components in our canal-ray tube where the field is applied at right angles to the beam.

A more extensive discussion of the whole complex phenomenon is for the time withheld since further consideration leads to some doubt as to whether the absorption acts precisely in the manner described in the paper just cited, and simple supplementary experiments at this laboratory have not disclosed the action with more certainty.

SUMMARY

The Stark effects for $H_\alpha D_\alpha$, $H_\beta D_\beta$, $H_\gamma D_\gamma$ and $H_\delta D_\delta$ have been photographed from a Lo Surdo source containing a mixture of the isotopes.

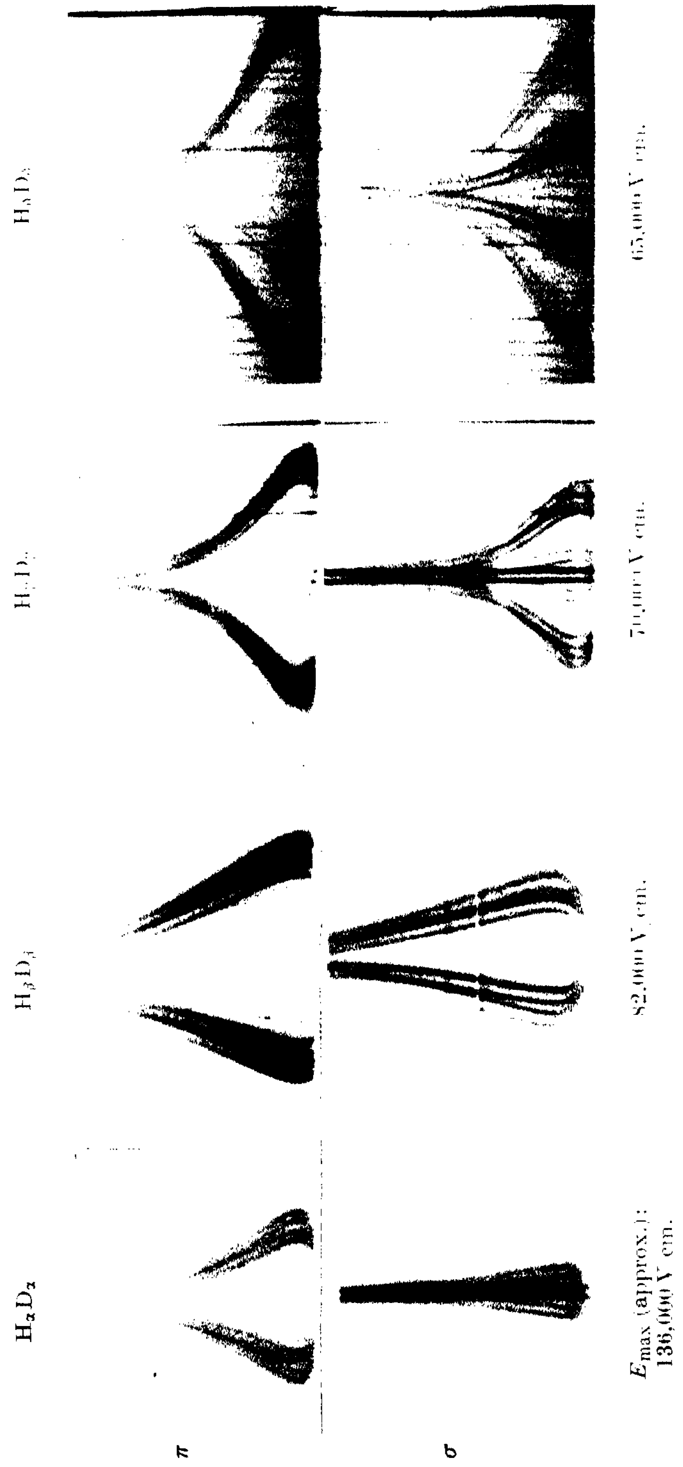
The observed asymmetric Stark intensity patterns for deuterium are nearly the mirror images of those for hydrogen. Somewhat similar results are obtained with a new canal-ray tube which operates at pressures much higher than usual.

These effects are attributed mainly to selective absorption with the source.

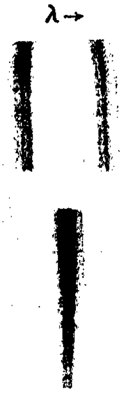
Variations from the Epstein spacing of components are found to be in fair agreement with the Schlapp theory of the fine structure of the Stark effect. The observations confirm earlier researches on hydrogen, and are extended to include deuterium.

REFERENCES

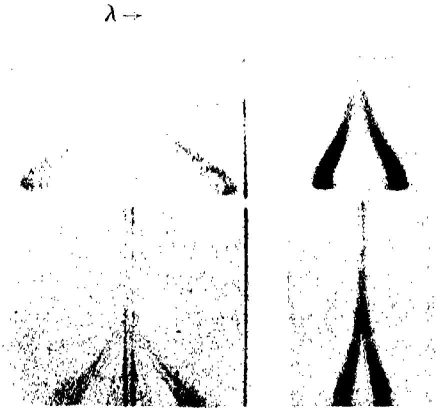
Foster, J. S. 1924 *Phys. Rev.* **23**, 667–84.
— 1925 *Astrophys. J.* **62**, 229–37.
— 1930 *J. Franklin Inst.* **209**, 585–624.
— 1934 *Nature*, **133**, 568.
Foster, J. S. and Chalk 1929 *Proc. Roy. Soc.* A, **123**, 108.
Foster, J. S. and Snell, H. 1934 *Nature, Lond.*, **153**, 558.
McRae, D. R. 1931 *Proc. Roy. Soc.* A, **132**, 257–65.
Schlapp, R. 1928 *Proc. Roy. Soc.* A, **119**, 313–34.
Schrödinger 1926 *Preuss. Akad. Wiss. Berlin, Ber.*, 23–26.
Thornton, R. L. 1935 *Proc. Roy. Soc.* A, **150**, 259–66.


$H_\alpha D_\alpha$
$H_\beta D_\beta$
$H_\gamma D_\gamma$
$H_\delta D_\delta$
π
σ
E_{max} (approx.): 136,000 V/cm.
82,000 V/cm.
70,000 V/cm.
65,000 V/cm.

(*Facing p.* 356)



(1)



(2)

The Influence of Hydrogen and Deuterium on the Thermal Decomposition of Diethyl Ether in the Low Pressure Region

By J. W. MITCHELL, *1851 Exhibition Scholar, New Zealand,*

AND C. N. HINSHELWOOD, F.R.S.

(*Received* 25 *June* 1937)

Richards (1936) found from sound dispersion measurements that at ordinary temperatures, in exciting the lower vibrational states of ethylene, collisions between hydrogen and ethylene molecules were ten times as effective, and collisions between deuterium and ethylene molecules 1·27 times as effective as collisions between ethylene molecules themselves. This means that hydrogen is about eight times as effective as deuterium. The specific effects of foreign gases in facilitating the interconversion of translational and vibrational energy as determined from measurements of the velocity of sound in the mixtures, have been explained by Franck and Eucken (1933) in terms of the chemical affinity between the colliding partners. This can however scarcely explain the results of Richards. The difference which he observed between hydrogen and deuterium seems to exclude any explanation for the superiority of hydrogen based on chemical affinity.

The specific effect of hydrogen in activating molecules in unimolecular decompositions has been described in previous communications from this laboratory (Hinshelwood 1927; Hinshelwood and Askey 1927; Hinshelwood and Staveley 1936). The comparison of hydrogen with helium seemed to suggest that the difference could not be explained by differences of mass and velocity but rather by differing chemical properties. It seemed that the comparison of the effect of hydrogen with that of deuterium would provide significant information, and that the comparison with the results of Richards would be of interest.

The unimolecular velocity constant for the thermal decomposition of diethyl ether begins to fall off in a convenient pressure range, and the effect of hydrogen in maintaining the unimolecular constant at low partial pressures of the ether is clearly shown. For this reason diethyl ether was chosen for the investigation. Staveley and Hinshelwood (1936) showed that in the ordinary thermal decomposition, chains of a mean length between three and four units were involved. To eliminate the influence of these chains, the experiments were made in the presence of 5 mm. of nitric oxide.

The essential result (Table III) is that, at 525° C., with the inhibited reaction in the low pressure region, deuterium has, within the limits of experimental error, no effect on the rate of reaction, whereas the addition of more than 400 mm. of hydrogen increases the rate to the value found with a high initial pressure of diethyl ether. This result provides definite evidence against any purely chemical explanation of the hydrogen effect.

An important factor governing specific energy transfers, is the mutual disturbance of the potential energy curves of the colliding molecules. Deuterium could not have an effect very different from hydrogen in this respect, as the electronic configurations of the two molecules are practically identical. The effect observed focuses attention primarily on the mass differences. This leads to a difference in the collision number of 1·4 times, and to a difference in the zero point energy which would correspond to a ratio of reaction rates of about two. Neither of these factors however nor a combination of both would account for the observations.

Nevertheless, it is possible with the aid of certain quite general assumptions about the conditions governing energy transfer to see how a small difference in mass could exert a very greatly magnified effect.

The necessary assumptions are not directly provable or disprovable but are not inherently improbable.

First it would have to be assumed that energy transfer only occurs during collisions, the duration of which is considerably greater than the average. This might arise from the fact that conditions were only favourable for energy transfer when the colliding molecules had reached a certain phase of vibration, the time required for this being in general greater than the collision time. If t is the time required for the suitable phase to be reached and θ the average duration of a collision, the probability of an encounter the duration of which exceeds t is given by the statistical formula $e^{-t/\theta}$. If t is increased to αt, the probability decreases in the ratio $e^{-(1-\alpha)t/\theta}$. When t/θ is great, the change in probability for a given value of α becomes very large. Now it might happen that t became smaller as the mass of one of the colliding partners was reduced. For example, if the frequency of the hydrogen or deuterium vibrations, or of the vibration of any bond involving a hydrogen or deuterium atom were concerned in the operations leading up to the energy transfer, then t would be smaller for hydrogen than for deuterium in a ratio calculable in principle from the masses. If this were the case, then the small difference of mass introduced into the exponential formula would have an effect increasing indefinitely as the ratio of t to θ

increased. For example, if the values of t varied as the square roots of the masses, a ratio of 5 to 1 for t/θ would account for the whole of the observed difference between hydrogen and deuterium. The mass function is doubtless much more complicated than this. Our only object is here to point out that a purely statistical factor might possibly magnify the mass differences enough to account for the observations without the introduction of hypotheses of an unusual or improbable nature.

At higher temperatures when the pressure of deuterium is greater than about 200 mm., there is a positive effect. This can be traced to the occurrence of an exchange reaction. 50 mm. of ether and 400 mm. of deuterium were allowed to react until no further pressure change occurred, a period of twelve hours. Assuming that there were 400 mm. of hydrogen plus deuterium present, and that the other gases were practically inert, the end products were used in making a fresh mixture of 50 mm. of ether and 200 mm. of hydrogen and deuterium. The initial rate observed when this reacted indicated the presence of 43 % hydrogen, since 200 mm. of pure deuterium had previously been shown to have little effect. To establish that even in the earlier stages of the reaction an exchange reaction occurred which would be capable of accounting for the apparent effect of deuterium at the higher temperature, samples of gas were taken during the course of the reaction. They were freed from gases other than hydrogen or deuterium by passage through a Pyrex trap immersed in liquid hydrogen, and then analysed by the microthermoconductivity method of Farkas. The results showed that during the reaction, 30–50 % of the deuterium was replaced by hydrogen.

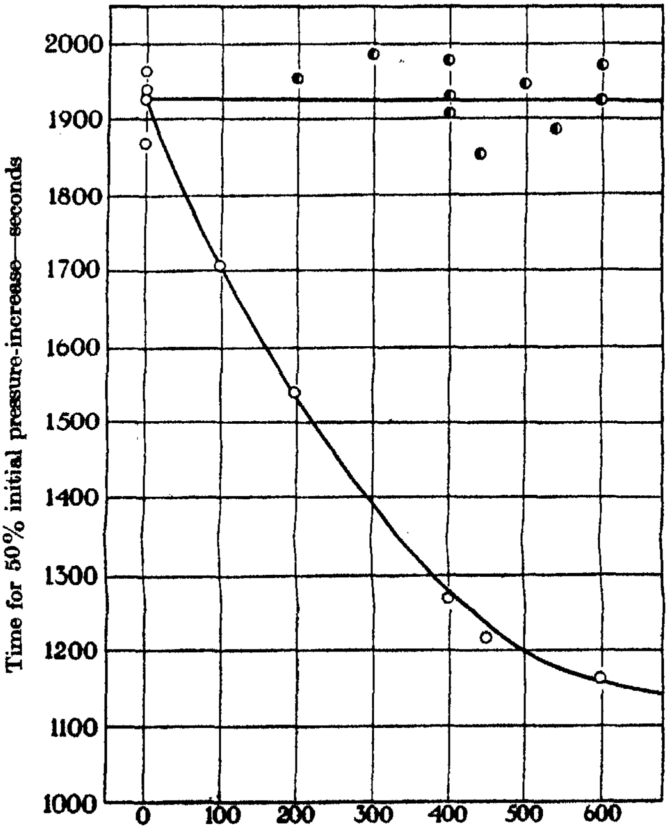
When experiments are made without the addition of nitric oxide, deuterium is found to have a very considerable effect though less than that of hydrogen. This is probably due to the interaction of the deuterium with the radicals which are present, giving rise to deuterium atoms by reactions such as the following:

$$R + \mathrm{D}_2 = R\mathrm{D} + \mathrm{D}.$$

Free atoms should be highly efficient in causing energy transfers, and as their mobility is greater than that of the methyl or ethyl radicals from which they arise, there could be a larger number of collisions in which they could exert their effect.

Besides exerting a direct catalytic effect, the deuterium atoms will certainly enter into exchange reactions leading to the formation of HD and eventually H_2 molecules.

A similar principle will account for the fact that hydrogen in the absence of nitric oxide is quantitatively more efficient in increasing the reaction rate than it is when the nitric oxide is present, the difference representing probably a contribution from the hydrogen atoms. These facts are reflected in a certain increase of the mean chain length with increase in the pressure of hydrogen or deuterium. This effect, however, diminishes with increasing pressure of the ether.



Pressure of added hydrogen or deuterium in mm.

Fig. 1—◐ deuterium; ○ hydrogen.

In connexion with this reaction $R + H_2 = RH + H$, the question arises whether the hydrogen atoms contributing to the subsequent decomposition of the ether can be as easily removed by nitric oxide as the radicals themselves. This can be tested by measuring the amount of nitric oxide required to reduce the decomposition rate to a given fraction of its original value. The measurements (Table VI) show that this amount is nearly the same whether hydrogen is present or not. From this we may conclude that nitric oxide is equally efficient as a chain breaker in reactions involving either organic radicals alone or radicals and hydrogen atoms.

This work having reopened the question of the influence of hydrogen on unimolecular reactions it seemed worth while to fill in a gap left in the earlier work, and to determine a temperature coefficient for the reaction in the presence of hydrogen. 50 mm. of ether were used and 300 mm. of hydrogen. The results given in Table VII lead to an activation energy of 60,000 cal. These measurements were not made with addition of nitric oxide, so that the activation energy is uncorrected for variation of chain length, and has not a direct physical significance. It is however directly comparable with the value found by Fletcher and Rollefson (1936) for the same ether pressure in the absence of hydrogen, namely 59,900. This supports other arguments in favour of the view that the action of the hydrogen is a physical one.

The experimental method was that normally used, but a few details call for mention. The temperature scale was established by direct calibration in terms of the rate of reaction of diethyl ether using previous results from this laboratory. The two extreme temperatures at which measurements were made correspond to times for a pressure increase of 50 % of the initial pressure of 93 and 323 sec. respectively. On this scale the maximum temperature deviation during measurements corresponds to time differences of ± 2 sec. and ± 8 sec. respectively.

Deuterium was prepared by electrolysis of heavy water of 99·6 % purity to which a small piece of metallic sodium was added, freed from oxygen by passing over a red hot tungsten spiral, and dried by passage through a trap immersed in liquid air. The hydrogen used was electrolytic hydrogen from a cylinder, purified and dried in the same way.

The numerical data upon which the preceding statements and discussion are based are collected in the following tables.

We are greatly indebted to Professor Simon and to Dr. Rollin for gifts of liquid hydrogen and for help with its manipulation.

SUMMARY

At 525° C. deuterium has been found to be less than one-tenth as efficient as hydrogen in activating molecules by collision in the chain-free decomposition of diethyl ether. At higher temperatures there is a positive effect but this can be traced to exchange reactions which give hydrogen.

The influence of hydrogen and of deuterium on the reaction in which the free radicals are not suppressed has also been further investigated.

The results are discussed.

Table I—The Variation of the Time for a Pressure Change of 50 % of the Initial Pressure to occur (t_{50}), with the Initial Pressure of Ether

Temperature 525° C.

Initial pressure of ether mm.	t_{50} sec.
487	262*
403	280*
400	263
382	306*
362	275
345	307
302	326*
300	307
	325
	323
217	337
210	333*
182	350
145	434
131	396
125	434
93	512*
77	537
50	550
	584
51	636
25	770
	765

Table I (*continued*)—The Effect of the Addition of approximately 300 mm. of Hydrogen on t_{50}

Temperature 525° C.

Pressure of ether mm.	Pressure of hydrogen mm.	t_{50} sec.
240	292	271*
206	293	280*
200	300	258
164	287	261*
154	300	257
100	312	237
75	300	250
69	292	287*
50	303	220
39	295	310*

* Results marked * are those of Hinshelwood (1927).

TABLE II—THE VARIATION OF t_{50} WITH THE PRESSURE OF ADDED HYDROGEN OR DEUTERIUM

Ether pressure 50 mm. throughout. Temperature 525° C.

Hydrogen		Deuterium	
Pressure mm.	t_{50} sec.	Pressure mm.	t_{50} sec.
0	584		
50	555	25	580
100	436	50	570
150	353	100	520
200	284	200	395
250	252	250	352
300	222	300	317
400	193		322
500	168	400	274
600	150	500	253
		600	250

TABLE III—THE VARIATION OF t_{50} WITH THE PRESSURE OF ADDED HYDROGEN OR DEUTERIUM, WHEN THE CHAINS HAVE BEEN COMPLETELY SUPPRESSED BY THE ADDITION OF 5 MM. OF NITRIC ACID

Temperature 525° C. Ether pressure 50 mm. throughout.

Hydrogen		Deuterium	
Pressure mm.	t_{50} sec.	Pressure mm.	t_{50} sec.
0	1940	200	1953
	1870	300	1985
100	1706	400	1930
200	1541	400	1973
300	1514	400	1908
400	1270	440	1855
450	1216	500	1945
600	1163	540	1885
		600	1970
		600	1923

Ether pressure 50 mm. Temperature 538° C.

Deuterium	
Pressure mm.	t_{25} sec.
0	442
	442
	466
453	487
522	447

Table IV—The Variation of the Apparent Chain Length with the Pressure of Hydrogen or Deuterium

Ether pressure 50 mm. Temperature 525° C.

Pressure of H_2 or D_2 mm.	Chain length	
	Hydrogen	Deuterium
0	3·25	3·25
100	3·90	3·64
200	5·44	4·80
300	6·30	5·94
400	6·60	6·95
500	7·25	7·46
600	7·59	7·55

Table V—The Variation of t_{50} with the Pressure of added Hydrogen, with and without the Addition of 5 mm. of Nitric Oxide

Temperature 556° C.

Pressure of ether mm.	Pressure of hydrogen mm.	t_{50} (0 mm. NO) sec.	t_{50} (5 mm. NO) sec.	Apparent chain length
50	0	144	486	3·38
50	100	—	390	—
50	200	90	334	3·71
50	300	59	304	5·15
50	400	55	276	5·01
50	500	50	262	5·26
300	0	93	—	—
		91	—	—
250	300	87	—	—
150	400	70	—	—

Table V (*continued*)—The Variation of t_{50} with the Pressure of added Deuterium, with and without the Addition of 5 mm. of Nitric Oxide

Temperature 556° C. Ether pressure 50 mm. throughout.

Pressure of deuterium mm.	t_{50} (0 mm. NO) sec.	t_{50} (5 mm. NO) sec.
0	139	486
	152	488
100	133	497
150	—	487
160	—	490
200	104	475

TABLE V—(*continued*)

Temperature 556° C. Ether pressure 50 mm. throughout.

Pressure of deuterium mm.	t_{50} (0 mm. NO) sec.	t_{50} (5 mm. NO) sec.
250	98	458
	—	456
	—	450
300	90	410
	—	415
350	—	390
400	86	385
	—	378
	—	375
	—	370
450	83	377
500	82	363

TABLE VI—VARIATION OF t_{50} WITH THE PRESSURE OF NITRIC OXIDE, WITH AND WITHOUT ADDITION OF HYDROGEN

Ether pressure 200 mm. throughout. Temperature 556° C.

Mean value of t_{50} for 200 mm. ether + 300 mm. hydrogen = 79 sec.: for 200 mm. ether alone = 104 sec. The values of the relative rates recorded below are reciprocals of t_{50} relative in each column to that for the experiment made without addition of nitric oxide, which is taken as 100.

Pressure of nitric oxide mm.	Relative rate without H_2	Relative rate with H_2
0	100	100
0·18	76	—
0·4	57	64
0·6	49	—
0·80	—	48
0·84	46	—
1·00	41	45
1·04	44	44
1·20	—	43·5
1·40	—	39
1·42	41	—
1·46	—	40
1·60	—	37
1·66	38	—
1·74	37	35
2·40	—	29
2·88	33	31
3·30	—	29
3·47	31	31
9·46	27	—
20·0	27	—

TABLE VII—ENERGY OF ACTIVATION FOR THE THERMAL DECOMPOSITION OF 50 MM. OF ETHER IN THE PRESENCE OF 300 MM. OF HYDROGEN

Temperature deg. abs.	t_{50} sec.
851	17
828	50
802	162
795	211
792	229
794	213
784	357
782	395
777	496
761	1144
738	4144

$E = 60{,}000$ cal.

REFERENCES

Fletcher and Rollefson 1936 *J. Amer. Chem. Soc.* **58**, 2129.
Franck and Eucken 1933 *Z. phys. Chem.* B, **20**, 460.
Hinshelwood 1927 *Proc. Roy. Soc.* A, **114**, 84.
Hinshelwood and Askey 1927 *Proc. Roy. Soc.* A, **115**, 215.
Hinshelwood and Staveley 1936 *J. Chem. Soc.* p. 818.
Richards 1936 *J. Chem. Phys.* **4**, 561.
Staveley and Hinshelwood 1936 *Proc. Roy. Soc.* A, **154**, 335.

Metallic Reflexion and the Surface Photoelectric Effect

By R. E. B. Makinson, B.Sc., *St. John's College, Cambridge*

(*Communicated by R. H. Fowler, F.R.S.—Received* 29 *May* 1937)

Introduction

1—The photoelectric effect at metal surfaces is of great importance and many attempts have been made to explain its main features, but an exact theory presents great difficulty. One of the best and most recent is that of Mitchell (1934, 1936), who considered an idealized model in which the potential barrier at the surface was a simple step; the field of the light wave was calculated by the classical optical theory, assuming that the optical constants change abruptly at the surface.

The latter approximation is not a good one, even for the simple potential jump, since the electron density does not fall sharply to zero, but decreases in a distance of about 3×10^{-8} cm. from a high value inside the metal to a negligible value outside the jump. It is well known that a free electron gives no photoelectric effect, and so the photo-emission from a metal is almost entirely due to the surface-potential gradient; it is in this region that the light wave is least accurately given by the simple optical theory, and fluctuates most.

Schiff and Thomas (1935) have given a quantum theory of metallic reflexion which enables the exact field of the light wave to be found, but their calculations are very complicated even in the simple case when the potential jump is infinite, and they did not try to find the photoelectric current quantitatively. The first object of the present paper is to give a semi-classical theory of metallic reflexion which is considerably simpler than that of Schiff and Thomas, while closely approximating to it. We use Sommerfeld's model of a metal, the potential energy of an electron being $-h\nu_a$ in the interior and changing abruptly to zero at the surface. With this model the electron density $n(x)$ is calculated as a function of x the distance from the surface, and has the form shown in fig. 1. We then calculate classically the effect of this electron density on the incident light wave.

This method differs in the following way from that of Schiff and Thomas. These authors find that part of the current which is coherent with an arbitrary light wave, whose scalar and vector potentials are Φ and $\mathbf{A}$, resulting from the perturbation of the wave function u of a single electron. This current density is given by

$$\mathbf{j} = \frac{eh}{4\pi mi}(u\nabla u^* - u^*\nabla u) - \frac{e^2}{mc}\mathbf{A}uu^*, \qquad (1)$$

where $-e$ is the charge and m the mass of the electron. The current densities due to all the conduction electrons are added together to form a resultant $\mathbf{J}$. When this is inserted into Maxwell's equations we have a set of integro-differential equations for Φ and $\mathbf{A}$. The present method amounts to using the unperturbed wave function in (1), which is equivalent to neglecting the first term; the problem is then much simpler. By this approximation we have excluded radiation damping (collision damping is neglected in both theories and is unimportant except in the far infra-red); the dielectric constant of the space near the surface, which of course varies with the co-ordinate x, then depends only on the electron density, and not on the individual electronic states. It is shown in § 4 that to the degree of approximation which Schiff and Thomas find manageable the two theories lead to the same expressions for the field, except at the point where the dielectric constant vanishes, the difference there being ascribed to the omission by these authors of an important term from the perturbed wave equation.

In § 5 an expression for the resulting photo-emission is found. It is shown that in Mitchell's formulae for the current in the case he considered a factor 2 is omitted, and that there is a further error through the omission of the above-mentioned term from the wave equation. The effect of roughness of the surface, on a scale small compared with the wave-length of the light, is briefly discussed in § 7. The results calculated for potassium are compared with experimental determinations of the current, and it is suggested that the poor agreement found may be due to the use of the simple step potential barrier in place of the type of barrier found by Bardeen (1936) for sodium, which is rounded and has an effective height depending on the momentum of the electron.

THE ELECTRON DENSITY NEAR THE SURFACE

2—The potential energy of a conduction electron in the metal is assumed to be

$$V = -h\nu_a, \quad x < 0,$$

$$= 0, \qquad x > 0.$$

Its wave function in the unperturbed state satisfies

$$\frac{h^2}{8\pi^2 m}\nabla^2\mu - \frac{h}{2\pi i}\frac{\partial u}{\partial t} - Vu = 0. \tag{2}$$

If
$$u = u_k = \psi_k \exp(-2\pi i E_k t/h), \tag{3}$$

then
$$\left.\begin{aligned}\psi_k &= \alpha_k(e^{-ik_1x} + a_k e^{ik_1x})\, e^{ik_2y+ik_3z}, & x<0,\\ &= \alpha_k b_k e^{-px} e^{ik_2y+ik_3z}, & x>0,\end{aligned}\right\} \tag{4}$$

with
$$p = (\mu\nu_a - k_1^2)^{\frac{1}{2}}, \quad \mu = 8\pi^2 m/h,$$

$$k_1^2 + k_2^2 + k_3^2 = \frac{8\pi^2 m}{h^2}(E_k + h\nu_a),$$

and a_k, b_k are determined by the continuity of ψ_k and $\partial\psi_k/\partial x$ at $x = 0$, giving

$$a_k = \frac{ik_1 - p}{ik_1 + p} = -e^{i\delta_k}, \quad \tan\delta_k = \frac{-2pk_1}{p^2 - k_1^2},$$

$$b_k = \frac{2ik_1}{p + ik_1}.$$

In (4), α_k is a normalizing factor for one electron per cell h^3 of phase space; k_1 may be restricted to positive values, since the inclusion of negative values gives the same states again, and merely alters the normalizing factor.

Thus in $x < 0$,
$$|\psi_k|^2 = 2|\alpha_k|^2[1 - \cos(2k_1x + \delta_k)],$$

giving inside the metal a mean partial electron density

$$2\overline{|\psi_k|^2} = 4|\alpha_k|^2,$$

which must be equated to

$$\frac{2\beta}{8\pi^3}\,\frac{dk_1\,dk_2\,dk_3}{1 + \exp\{h^2(k_1^2 + k_2^2 + k_3^2 - k_0^2)/8\pi^2 mkT\}}.$$

Here k is Boltzmann's constant, T is the absolute temperature, k_0 corresponds to the limit of the Fermi distribution of momentum, and β will be determined from the relation

$$k_0 = (3\pi^2 n_0)^{\frac{1}{3}}, \tag{5}$$

n_0 being the number of conduction electrons per unit volume in the interior of the metal. Thus when $T = 0$, and also approximately at ordinary temperatures,

$$|\alpha_k|^2 = \frac{\beta}{16\pi^3}\,dk_1\,dk_2\,dk_3.$$

Hence, integrating over all possible values of **k**, that is over the hemisphere in **k** space of radius k_0 lying in $k_1 > 0$, the total density is, in $x < 0$,

$$n(x) = 2\sum_k |\psi_k|^2 = \frac{\beta}{4\pi^2}\int_0^{k_0}[1-\cos(2k_1x+\delta_k)]\,(k_0^2-k_1^2)\,dk_1, \tag{6}$$

and similarly in $x > 0$

$$n(x) = \frac{\beta}{2\pi^2\mu\nu_a}\int_0^{k_0} e^{-2\mu x}k_1^2(k_0^2-k_1^2)\,dk_1. \tag{6a}$$

From (6) we find for the density in the interior of the metal

$$n_0 = \frac{\beta k_0^3}{6\pi^2},$$

thus, by (5), $\beta = 2$.

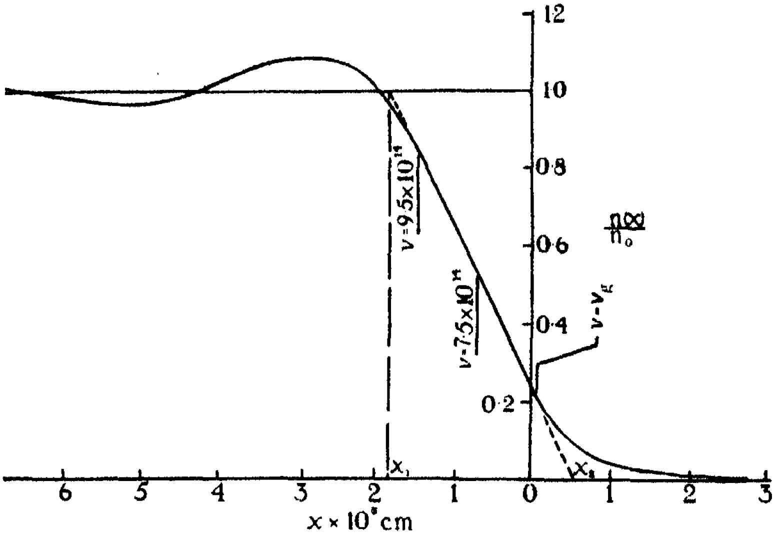


FIG. 1—Density of conduction electrons near the surface of potassium, assuming a square potential barrier. The point at which the dielectric constant vanishes is shown for three frequencies.

For potassium, assuming the values

$$n_0 = 1{\cdot}35 \times 10^{22},$$
$$k_0 = 7{\cdot}37 \times 10^{7},$$
$$\nu_a = 9{\cdot}8 \times 10^{14},$$

the resulting density is shown in fig. 1. It is seen that $n(x)$ becomes constant at a distance inside the surface very much less than the wave-length of light in the neighbourhood of the threshold.

THE FIELD OF THE LIGHT WAVE

3—Under the influence of a light wave the electrons oscillate, and their density becomes say $n(x)+n'(x,y,z,t)$. Neglecting damping, we have to the first order as the equation of motion of electrons with charge $-e$ and velocity $\mathbf{v}$,

$$\frac{d\mathbf{v}}{dt} = \frac{-\mathbf{E}e}{m}.$$

Also

$$\left.\begin{aligned} \operatorname{curl}\mathbf{E} &= -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \\ \operatorname{curl}\mathbf{H} &= \frac{1}{c}\left(\frac{\partial \mathbf{E}}{\partial t} - 4\pi n e\mathbf{v}\right), \\ \operatorname{div}\mathbf{E} &= -4\pi n' e, \\ \operatorname{div}\mathbf{H} &= 0. \end{aligned}\right\} \tag{7}$$

If $\mathbf{E}$ and $\mathbf{H}$ have frequency ν, the third equation becomes

$$\operatorname{curl}\mathbf{H} = \frac{1}{c}\frac{\partial(\epsilon\mathbf{E})}{\partial t},$$

where

$$\epsilon = 1 - \frac{n(x)\,e^2}{\pi m \nu^2}, \tag{8}$$

and since

$$\frac{\partial n'}{\partial t} = -n\operatorname{div}\mathbf{v} - \mathbf{v}\cdot\operatorname{grad} n,$$

the fourth equation of (7) becomes

$$\operatorname{div}(\epsilon\mathbf{E}) = 0. \tag{9}$$

The equations for $\mathbf{E}$ and $\mathbf{H}$ are thus those for a medium of dielectric constant ϵ. Since ϵ is a function of x only (fig. 2), and $\partial/\partial z = 0$, we have (Foersterling 1928, p. 277)

$$\frac{\epsilon}{c}\frac{\partial E_x}{\partial t} = \frac{\partial H_z}{\partial y}, \tag{10}$$

$$\frac{\epsilon}{c}\frac{\partial E_y}{\partial t} = -\frac{\partial H_z}{\partial x}, \tag{11}$$

$$\frac{1}{c}\frac{\partial H_y}{\partial t} = \frac{\partial E_z}{\partial x}. \tag{12}$$

Hence
$$\frac{\partial^2 H_z}{\partial x^2} - \frac{1}{\epsilon}\frac{\partial \epsilon}{\partial x}\frac{\partial H_z}{\partial x} + \frac{\partial^2 H_z}{\partial y^2} + \frac{4\pi^2\nu^2}{c^2}\epsilon H_z = 0, \tag{13}$$

$$\frac{\partial^2 H_x}{\partial x^2} + \frac{\partial^2 H_x}{\partial y^2} + \frac{4\pi^2\nu^2}{c^2}\epsilon H_x = 0, \tag{14}$$

$$\frac{\partial^2 E_z}{\partial x^2} + \frac{\partial^2 E_z}{\partial y^2} + \frac{4\pi^2\nu^2}{c^2}\epsilon E_z = 0. \tag{15}$$

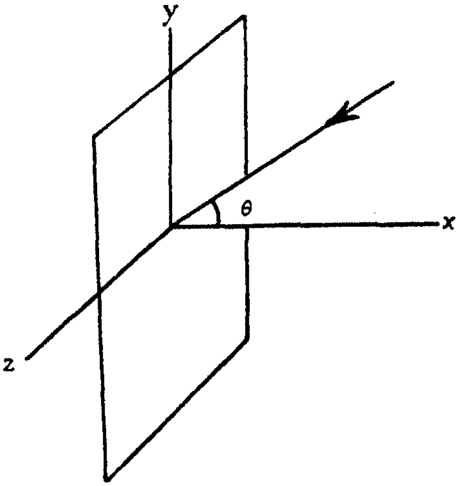


Fig. 2

Incident Light Polarized in the Plane of Incidence

3·1—If the angle of incidence is θ (fig. 2), (13) becomes

$$\epsilon(x)\frac{\partial^2 H_z}{\partial x^2} - \frac{\partial \epsilon(x)}{\partial x}\frac{\partial H_z}{\partial x} + \frac{4\pi^2\nu^2\epsilon(x)}{c^2}[\epsilon(x) - \sin^2\theta]H_z = 0, \tag{16}$$

since
$$\frac{\partial^2 H_z}{\partial y^2} = -\frac{4\pi^2\nu^2}{c^2}\sin^2\theta\, H_z.$$

The light will thus be totally reflected if

$$1 - \frac{n_0 e^2}{\pi m \nu^2} = \epsilon_0 < \sin^2\theta, \tag{17}$$

that is if
$$\sin^2\theta - \epsilon_0 = \kappa^2 > 0, \tag{18}$$

ϵ_0 being the steady value ϵ takes in the interior of the metal, and κ the optical extinction constant of the Drude theory. The wave system near the surface is then stationary, the phase being independent of x. In potassium the inequality (17) is satisfied for frequencies up to $1{\cdot}04 \times 10^{15}$ for normal incidence, and up to $1{\cdot}62 \times 10^{15}$ when $\theta = 50^\circ$; above these frequencies the metal is transparent. (The experimental value for the frequency at which the metal becomes transparent is $0{\cdot}95 \times 10^{15}$ when $\theta = 0$ (Wood 1933).) The high reflecting power found at lower frequencies (Frehafer 1920) shows that the neglect of damping gives a good approximation. We assume in the following that (17) is satisfied, since this covers the frequencies and angles at which we are interested in the photoelectric effect.

To solve (16), put

$$H_z = De^{2\pi\nu\kappa x/c}[1+f(x)], \quad x<0,$$

where $f(-\infty) = f'(-\infty) = 0$, and we have omitted a factor

$$\mathscr{R} \exp[-2\pi i\nu(t+y\sin\theta/c)],$$

$\mathscr{R}$ denoting the real part. Then

$$\epsilon f'' + \left(\frac{4\pi\nu\epsilon\kappa}{c} - \epsilon'\right)f' + \left\{\frac{4\pi^2\nu^2\epsilon}{c^2}(\epsilon-\epsilon_0) - \frac{2\pi\nu\epsilon'\kappa}{c}\right\}f = \frac{2\pi\nu\epsilon'\kappa}{c} - \frac{4\pi^2\nu^2\epsilon}{c^2}(\epsilon-\epsilon_0), \tag{19}$$

where primes denote differentiation with respect to x. Since the variations in $\epsilon(x)$ occupy a length much less than a light wave-length, we may neglect $2\pi\nu(\epsilon-\epsilon_0)/c$ in comparison with $\kappa\epsilon'$. An approximate solution satisfying the boundary conditions is then

$$f = \frac{2\pi\nu\kappa}{c\epsilon_0}\int_{-\infty}^{x}(\epsilon-\epsilon_0)\,dx,$$

since on substitution in (19), all but negligible terms disappear.

Thus when $x<0$

$$H_z = De^{2\pi\nu\kappa x/c}\left[1+\frac{2\pi\nu\kappa}{c\epsilon_0}\int_{-\infty}^{x}(\epsilon-\epsilon_0)\,dx\right], \tag{20}$$

and similarly we find approximately for $x>0$

$$H_z = \left[A - B\frac{2\pi\nu\cos\theta}{c}\int_{\infty}^{x}(1-\epsilon)\,dx\right]\cos\frac{2\pi\nu x\cos\theta}{c} + \left[B + A\frac{2\pi\nu\cos\theta}{c}\int_{\infty}^{x}(1-\epsilon)\,dx\right]\sin\frac{2\pi\nu x\cos\theta}{c}. \tag{21}$$

The constants A, B, D, may be related by joining these solutions in magnitude and gradient at the point near x_1 (fig. 1) where $n = n_0$, that is, $\epsilon = \epsilon_0$. The second terms in each square bracket of (20) and (21) are small compared with the first unless ϵ_0 is very nearly equal to zero or to $\sin^2\theta$, though their derivatives are appreciable. Without these terms H_z has the form appropriate to a sharp discontinuity in ϵ from ϵ_0 to unity, and with them it has still very nearly a constant value $H_z(0)$ over many electron wave-lengths from the surface. From the joining conditions we find

$$\left.\begin{aligned} A &= D = H_z(0), \\ B &= \frac{\kappa}{\epsilon_0 \cos\theta}. \end{aligned}\right\} \tag{22}$$

We now find E_x and E_y from (10) and (11):

$$E_x = \frac{\sin\theta}{\epsilon(x)} H_z, \tag{23}$$

$$E_y = \frac{-c}{\epsilon(x)} \frac{\partial H_z}{\partial x}. \tag{24}$$

In (23) we may put $H_z = H_z(0)$ at points much less than a light wave-length from the surface. We see that E_x becomes infinite at the point x_0 where ϵ vanishes. The roots of the indicial equation of (16) corresponding to solutions in powers of $\xi = x - x_0$ are 0 and 2, and the solution for H_z contains a term of the form $\xi^2 \log \xi$. Thus from (24) and (16) E_y is logarithmically infinite at $x = x_0$. If we take account of damping by putting

$$\epsilon(x) = \epsilon_1(x) + i\epsilon_2, \tag{25}$$

then E_x and E_y remain finite at x_0. In the following we suppose that ϵ_2 is very small (it is necessarily positive), and that ϵ_1 is given by (8).

From (21), if $H_z(\infty)$ is the amplitude of H_z in the incident *progressive* wave,

$$H_z(0) = \frac{-2\epsilon_0 \cos\theta}{(1-\epsilon_0)^{\frac{1}{2}}\{1-(1+\epsilon_0)\cos^2\theta\}^{\frac{1}{2}}} H_z(\infty) \tag{26}$$

approximately, which is the same relation as for a sharp discontinuity in ϵ. The rate of incidence of energy on the surface is thus

$$\frac{c}{32\pi \times 4{\cdot}18 \times 10^7} \frac{(1-\epsilon_0)\{1-(1+\epsilon_0)\cos^2\theta\}}{\epsilon_0^2 \cos\theta} H_z(0)^2 \text{ calories cm.}^{-2}\text{ sec.}^{-1}. \tag{27}$$

3·2—In calculating the photo-emission we require to know the vector and scalar potentials **A** and Φ, also div **A**, near the surface. These are not uniquely determined by **E** and **H**; we may impose the condition that everywhere

$$\Phi = 0, \tag{28}$$

or that

$$\operatorname{div}\mathbf{A} + \frac{1}{c}\frac{\partial\Phi}{\partial t} = 0, \tag{29}$$

but not in general both. If we take the condition (28), div **A** is not zero (except where $\partial\epsilon/\partial x = 0$), for this requires $\operatorname{div}\mathbf{E} = 0$, whereas $\operatorname{div}(\epsilon\mathbf{E}) = 0$. In fact the div **A** so obtained becomes very large near points where ϵ vanishes or changes abruptly.

We suppose that (28) is the condition satisfied, since the results are then expressible more simply. Let

$$\mathbf{A}(x) = \mathbf{a}(x)\exp\{-2\pi i\nu(y\sin\theta/c + t)\} + \text{conjugate};$$

then from (23) and (24), since

$$\frac{\partial\mathbf{A}}{\partial t} = -c\mathbf{E},$$

we have, at points whose distance from the surface is much less than a light wave-length,

$$\left.\begin{aligned} a_x &= \frac{-ic}{4\pi\nu}\frac{1}{\epsilon}H_z(0)\sin\theta, \\ a_y &= \frac{-c^2}{8\pi^2\nu^2}\frac{1}{\epsilon}\frac{\partial H_z}{\partial x}, \\ a_z &= 0, \end{aligned}\right\} \tag{30}$$

near the surface. Also

$$\operatorname{div}\frac{\partial\mathbf{A}}{\partial t} = -c\operatorname{div}\mathbf{E},$$

and, by (9),

$$\operatorname{div}\mathbf{E} = -\frac{1}{\epsilon}\frac{\partial\epsilon}{\partial x}E_x,$$

thus, by (23),

$$\operatorname{div}\mathbf{a} = \frac{ic}{4\pi\nu}\frac{1}{\epsilon^2}\frac{\partial\epsilon}{\partial x}H_z(0)\sin\theta. \tag{31}$$

Incident Light Polarized Perpendicular to the Plane of Incidence

3·3—From the equations (13), (10) and (11), $H_z = E_x = E_y = 0$ everywhere, and

$$\frac{\partial^2 E_z}{\partial x^2} - \frac{4\pi^2\nu^2}{c^2}(\sin^2\theta - \epsilon)E_z = 0. \tag{32}$$

The wave system is stationary as before, if (17) is satisfied. Let

$$E_z = Ce^{2\pi\nu\kappa x/c}\{1+g(x)\}, \quad x<0,$$

where $g(-\infty) = g'(-\infty) = 0$, and a factor $\mathscr{R}\exp\{-2\pi i\nu(t+y\sin\theta/c)\}$ is omitted. Then over the region where ϵ is sensibly different from ϵ_0, and g from zero,

$$g'' - \frac{4\pi^2\nu^2}{c^2}(\sin^2\theta-\epsilon)g = \frac{4\pi^2\nu^2}{c^2}(\epsilon_0-\epsilon),$$

so, neglecting the second term on the left,

$$E_z = Ce^{2\pi\nu\kappa x/c}\left\{1+\frac{4\pi^2\nu^2}{c^2}\int_{-\infty}^{x}\int_{-\infty}^{x}(\epsilon_0-\epsilon)\,dx\,dx\right\}. \tag{33}$$

We see from (33) that E_z and $\partial E_z/\partial x$ are very little affected by the fluctuation in $n(x)$, and vary slowly near the surface.

Joining the solution (33) to that outside the metal we have, if $E_z(\infty)$ is the amplitude of the incident progressive wave,

$$E_z(\infty) = \frac{(1-\epsilon_0)^{\frac{1}{2}}}{2\cos\theta}E_z(0), \tag{34}$$

which is the same relation as for a sharp discontinuity in ϵ. The rate of incidence of energy per cm.^{2} of surface is thus, in calories per second,

$$\frac{c}{32\pi\times 4{\cdot}18\times 10^7}\,\frac{(1-\epsilon_0)\,E_z(0)^2}{\cos\theta}. \tag{35}$$

In this case, since $H_z = 0$, div **A** is by (31) everywhere zero.

COMPARISON WITH THE METHOD OF SCHIFF AND THOMAS

4—When a light wave acts on the electrons, the wave equation (e.g. Kronig and Groenewold 1934), omitting a term in $\mathbf{A}^2$, is

$$\frac{h^2}{8\pi^2m}\nabla^2u-\frac{h}{2\pi i}\frac{\partial u}{\partial t}-Vu = \frac{-ihe}{2\pi mc}(\mathbf{A}\cdot\operatorname{grad}u+\tfrac{1}{2}u\operatorname{div}\mathbf{A})-e\Phi u. \tag{36}$$

Let $u = u_k+v_k$, where u_k is the solution (3) of the unperturbed wave equation (2). Then to the first order

$$\frac{h^2}{8\pi^2m}\nabla^2v_k-\frac{h}{2\pi i}\frac{\partial v_k}{\partial t}-Vv_k = \frac{-ihe}{2\pi mc}(\mathbf{A}\cdot\operatorname{grad}u_k+\tfrac{1}{2}u_k\operatorname{div}\mathbf{A})-e\Phi u_k. \tag{37}$$

Schiff and Thomas have taken $\Phi = 0$ for all x, but as we have seen in § 3·2 we cannot omit the term in div **A** in (36) as these authors have done.

The current coherent with the light wave is given for the pair† of electrons with wave function u by

$$d\mathbf{J} = \frac{2eh}{4\pi mi}(u\nabla u^* - u^*\nabla u) - \frac{2e^2}{mc}\mathbf{A}uu^*. \tag{38}$$

This is to be integrated over all the conduction electrons to give the total current **J**. Then from Maxwell's equations

$$\begin{aligned}\operatorname{curl}\operatorname{curl}\mathbf{A} + \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} &= \frac{4\pi}{c}\mathbf{J}\\ &= \frac{ieh}{4\pi^2mc}\int_0^{k_0}(k_0^2-k_1^2)\,(u_k'^*\,\nabla v_k' - v_k'\,\nabla u_k'^* + v_k'^*\,\nabla u_k' - u_k'\,\nabla v_k'^*)\,dk_1\\ &\qquad - \frac{e^2}{\pi mc^2}\mathbf{A}\int_0^{k_0} u_k' u_k'^*(k_0^2-k_1^2)\,dk_1,\end{aligned} \tag{39}$$

and this integro-differential equation has to be solved for **A**. After some approximation Schiff and Thomas reduce (39) to the form

$$B_x(x) = \frac{\chi}{h(x) + \int_{-\infty}^{0}[K(x,s)\,B_x(s)/B_x(x)]\,ds}, \tag{40}$$

where $\quad \mathbf{A} = B(x)\exp\{-2\pi i\nu(t - y\sin\theta/c)\} + \text{conjugate}.$

The equation corresponding to (39) on the basis of § 3, when $\Phi = 0$, is

$$\begin{aligned}\operatorname{curl}\operatorname{curl}\mathbf{A} + \frac{1}{c^2}\frac{\partial^2\mathbf{A}}{\partial t^2} &= \frac{-4\pi e^2}{mc^2}n(x)\,\mathbf{A}\\ &= \frac{-e^2}{\pi mc^2}\mathbf{A}\int_0^{k_0} u_k' u_k'^*(k_0^2-k_1^2)\,dk_1.\end{aligned} \tag{41}$$

Thus any difference between the results of the two methods is due to the first term on the right side of (39), which does not appear in (41) because no account is there taken of the perturbed form of the wave functions, the Fermi distribution of momentum being supposed unchanged when the electrons oscillate. The term in question in (39) gives the effect of radiation damping, which we have thus neglected. This term leads to the integral in the denominator in (40), which Schiff and Thomas show to be negligible except as to the sign of its imaginary part at the zero of $h(x)$. They find that the sign changes at $\nu = 6\cdot 13\times 10^{14}$ in the case of potassium; since, however, they have omitted the term in div **A** from (36), this needs re-examination.

---

† The fact that the two electrons of opposite spin corresponding to each eigenfunction u both contribute to the current seems to have been lost sight of in some previous work; thus the right side of Mitchell's (1934) equations (33), (35) should be doubled.

Equation (40) is equivalent to our (23), except near the zero of $h(x)$, for it may be shown that

$$h(x) = \frac{4\pi^2\nu^2}{c^2}\epsilon_1(x)$$

in the case they have considered with the potential barrier at the surface an infinite jump; and also that

$$\chi = \frac{i\pi\nu}{c}H_z(0)\sin\theta$$

in terms of the present notation. Corresponding to the imaginary part in the denominator of (40) is the imaginary part ϵ_2 of ϵ in the denominator of (23) arising from slight damping. The change of sign which Schiff and Thomas deduced would mean negative damping at some frequencies.

Thus we see that the present method of calculating the field gives results closely approximating to those of the more accurate quantum theory of reflexion.

The Photoelectric Current

5—We gave in § 4 the first-order equation (37) for the perturbation produced by the light wave, and in this we put $\Phi = 0$ in accordance with (28). Mitchell (1934, equations (12), (57), (61)) has found by the method of variation of constants the solution v_k for x large and positive, when stimulated emission is neglected, of the equation

$$\frac{h^2}{8\pi^2 m}\nabla^2 v_k + \frac{h}{2\pi i}\frac{\partial v_k}{\partial t} - Vv_k = \frac{-ihe}{2\pi mc}\mathbf{A}\cdot\operatorname{grad} u_k, \tag{42}$$

with $$u_k = \psi_k \exp(+2\pi i E_k t/h),$$

from which (37) differs in that on the right side

$$\left.\begin{array}{ll} & \mathbf{A}\cdot\operatorname{grad} u_k + \tfrac{1}{2}u_k \operatorname{div}\mathbf{A} \\ \text{replaces} & \mathbf{A}\cdot\operatorname{grad} u_k, \end{array}\right\} \tag{43}$$

and the sign of the second term of (42) is changed to the more usual minus. It is easily verified that we may use Mitchell's solution of (42) for v_k outside the metal, on making in it the substitution (43). This solution is

$$v_k = \frac{4\pi^2 m}{hr} b_r (k\,|\,\mathbf{a}\,|\,r)_0^+ e^{-irx+ik_2y+ik_3z} e^{+2\pi i(E_k+h\nu)t/h}, \tag{44}$$

where
$$b_r = \frac{2q}{q+r},$$

$$q = (\mu\nu + k_1^2)^{\frac{1}{2}},$$

$$r = \{k_1^2 + \mu(\nu - \nu_a)\}^{\frac{1}{2}},$$

and (with a slight change of notation)

$$(k\,|\,\mathbf{a}\,|\,r)_0^+ = \frac{ie}{\pi mc}\frac{\alpha_k r k_1}{q(p+ik_1)} G_0(\mathbf{k}). \tag{45}$$

Here α_k and p are the same as in § 2, and

$$G_0(\mathbf{k}) = \int_{-\infty}^{\infty} \left\{\frac{d\chi_k}{dx} a_x^* + \chi_k(ik_2 a_y^* + ik_3 a_z^*)\right\} \chi_r^*\, dx,$$

in which
$$\left.\begin{aligned} \chi_k &= \cos k_1 x - \frac{p}{k_1}\sin k_1 x, && x < 0,\\ &= e^{-\nu x}, && x > 0, \end{aligned}\right\} \tag{46}$$

$$\left.\begin{aligned} \chi_r^* &= e^{iqx} && x < 0,\\ &= \frac{iq}{r}\sin rx + \cos rx, && x > 0, \end{aligned}\right\} \tag{47}$$

and a convergence factor $e^{\gamma x}$, where $0 < \gamma \ll k_1$, is used in $x < 0$. Making the substitution (43), we replace $G_0(\mathbf{k})$ in (45) by

$$G(\mathbf{k}) = -\int_{-\infty}^{\infty} \left\{\frac{d\chi_k}{dx} a_x + \chi_k(ik_2 a_y + ik_3 a_z + \tfrac{1}{2}\operatorname{div}\mathbf{a})\right\} \chi_r\, dx, \tag{48}$$

and (44) becomes

$$v_k = \frac{4\pi^2 m}{hr} b_r (k\,|\,\mathbf{a}\,|\,r)_0^+ e^{irx+ik_2y+ik_3z} e^{-2\pi i(E_k+h\nu)t/h}.$$

The contribution of the two† electrons with wave functions u to the current density outside the metal is given by (38), so to the first order

$$\begin{aligned} dJ_x &= \frac{2eh}{4\pi mi}\left(v_k\frac{\partial v_k{}^*}{\partial x} - v_k{}^*\frac{\partial v_k}{\partial x}\right) \\ &= -\frac{16\pi^3 me}{hr} b_r^2 \,|\, (k\,|\,\mathbf{a}\,|\,r)_0^+ \,|^2 \\ &= -\frac{e^3}{\pi^4 m^2 c^2 \nu_a}\frac{rk_1^2}{(q+r)^2}\,|G(\mathbf{k})|^2\, dk_1\, dk_2\, dk_3, \qquad r^2 > 0, \\ &= 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad r^2 < 0. \end{aligned}$$

This is to be integrated over the segment of the sphere in $\mathbf{k}$ space:

$$\begin{aligned} k_1^2 + k_2^2 + k_3^2 &\leqslant k_0^2, \\ k_1 &> 0, \\ r^2 &> 0, \end{aligned}$$

giving the photoelectric current in e.s.u. per cm.^{2}

$$J_x = -\frac{e^3}{\pi^4 m^2 c^2 \nu_a}\iiint \frac{rk_1^2}{(q+r)^2}\,|\, G(\mathbf{k}) \,|^2\, dk_1\, dk_2\, dk_3. \tag{49}$$

The yield in coulombs per calorie is then given by combining (49) and (27) or (35) according to the state of polarization of the incident light.

CALCULATIONS FOR POTASSIUM

6—To simplify the integrations in (48), we suppose $n(x)$ to be given by the broken line in fig. 1, so that

$$\left.\begin{aligned} \epsilon_1 &= \epsilon_0, & x &< x_1, \\ &= \epsilon_0 + \frac{(1-\epsilon_0)(x-x_1)}{x_2 - x_1}, & x_1 &< x < x_2, \\ &= 0, & x &> x_2. \end{aligned}\right\} \tag{50}$$

The values of x at which $\epsilon = 0$ are there shown for the threshold frequency $\nu_g = 4{\cdot}84 \times 10^{14}$, and for two higher frequencies. It is clear that the difference between this approximate variation of n and the more accurate one near the tail of the curve will not be important at frequencies above the threshold.

---

† See footnote, § 4.

There will be some error in the immediate neighbourhood of the frequency $\nu_0 = 1{\cdot}04 \times 10^{15}$ at which $\epsilon_0 = 0$, but many of the relations of § 3 are not then valid in any case.

Polarization in the Plane of Incidence

6·1—From (48), (30) and (31)

$$G(\mathbf{k}) = \frac{icH_z(0)\sin\theta}{4\pi\nu}\int_{-\infty}^{\infty}\left\{\chi_r\frac{d\chi_k}{dx}\frac{1}{\epsilon}+\tfrac{1}{2}\chi_r\chi_k\frac{\partial\epsilon^{-1}}{\partial x}\right\}dx - ik_2\int_{-\infty}^{\infty}\chi_r\chi_k a_y\,dx. \tag{51}$$

The last term vanishes if a_y is constant near the surface. From (30), when $x < x_1$ or $x > x_2$,

$$a_y = \frac{-c\kappa}{4\pi\nu\epsilon_0}H_z(0).$$

Since
$$\operatorname{curl}\mathbf{A} = \mathbf{H},$$

we have
$$\frac{\partial A_y}{\partial x} = \frac{\partial A_x}{\partial y} + H_z,$$

that is
$$\frac{\partial a_y}{\partial x} = \frac{1}{2}\left\{1-\frac{\sin^2\theta}{\epsilon(x)}\right\}H_z(0). \tag{52}$$

The last term of (51) is thus

$$-\frac{ik_2H_z(0)}{2}\int_{x_1}^{x_2}\chi_r\chi_k\int_{x_1}^{x}\left\{1-\frac{\sin^2\theta}{\epsilon(x)}\right\}dx\,dx,$$

which is found to make a negligible contribution to $G(\mathbf{k})$; so, writing

$$\frac{8\pi\nu\,G(\mathbf{k})}{icH_z(0)\sin\theta} = L(k_1),$$

we have

$$L(k_1) = 2\int_{-\infty}^{\infty}\left\{\chi_r\frac{d\chi_k}{dx}\frac{1}{\epsilon}+\tfrac{1}{2}\chi_r\chi_k\frac{\partial\epsilon^{-1}}{\partial x}\right\}dx = \int_{-\infty}^{\infty}\frac{1}{\epsilon_1+i\epsilon_2}\left\{\chi_r\frac{d\chi_k}{dx}-\chi_k\frac{d\chi_r}{dx}\right\}dx, \tag{53}$$

on integration by parts. If $\nu < \nu_0$, the integrand has a pole at the point x_0 where ϵ_1 vanishes, and a residue term contributes to $L(k_1)$. The integral over the ranges $x < x_1$ and $x > x_2$ is easily expressed exactly, and over the range

$x_1 < x < x_2$ it may be evaluated with the aid of cosine- and sine-integral tables, or by putting the integrand into the form

$$\frac{\chi(x_0)}{\epsilon_1 + i\epsilon_2} + \frac{\chi(x) - \chi(x_0)}{\epsilon_1 + i\epsilon_2},$$

where $$\chi = \chi_r \frac{d\chi_k}{dx} - \chi_k \frac{d\chi_r}{dx},$$

and integrating the second expression graphically between the limits x_1, x_2.

When (53) is evaluated, using the expressions (46) and (47) for χ_k and χ_r and (25), (50) for ϵ, we have from (49)

$$J_x = \frac{-e^3 \sin^2\theta\, H_z(0)^2}{64\pi^5 m^2 \nu_a \nu^2} \int_{0,[\mu(\nu_a - \nu)]^{\frac{1}{2}}}^{k_0} \frac{r k_1^2 (k_0^2 - k_1^2)}{(q+r)^2} \,|\, L(k_1) \,|^2 \, dk_1, \tag{54}$$

the zero limit of integration being taken if $\nu > \nu_a$. From (27) and (54), on giving the constants their numerical values, the following expression is obtained for the photoelectric yield from potassium in coulombs per calorie, when the polarization is in the plane of incidence:

$$P_{\parallel} = 3{\cdot}28 \times 10^{-4} \frac{\sin^2\theta \cos\theta\, \epsilon_0^2}{\nu^2 (1 - \epsilon_0) \{1 - (1 + \epsilon_0) \cos^2\theta\}}$$
$$\times \int_{0,[\mu(\nu_a - \nu)]^{\frac{1}{2}}}^{k_0} dk_1 \frac{r k_1^2 (k_0^2 - k_1^2)}{(q+r)^3} \,|L(k_1)\,|^2. \tag{55}$$

This expression for $P_{\parallel}$, when $\theta = 50°$, is plotted in fig. 3 against $(\nu - \nu_g)/\nu_g$, ν_g being the threshold frequency. It is to be noted that (55) is only valid when the light waves form a stationary system at the metal surface, that is, when the inequality (17) is satisfied. The derivation also breaks down in the immediate neighbourhood of the frequency ν_0 at which $\epsilon_0 = 0$.

Polarization Perpendicular to the Plane of Incidence

6·2—We have in this case $A_x = A_y = \operatorname{div} \mathbf{A} = 0$, and A_z is given by (33). If E_z were strictly constant near the surface, $G(\mathbf{k})$ would vanish identically; actually the difference of (33) from a constant value is so small over many electron wave-lengths from the surface that the resulting yield $P_{\perp}$ is estimated to be of the order 10^{-12} coulombs per calorie, which is quite negligible.

Comparison with Mitchell's Results

6·3—On reducing to zero width the region in which ϵ changes from unity to ϵ_0, we obtain from (53) and (54) the photoelectric current in the case which Mitchell (1934, 1936) considered. In (50) let x_1 and x_2 approach zero; we then find from (53)

$$L(k_1) = -\frac{2(1-\epsilon_0)}{\mu\nu\epsilon_0}(ipq+k_1^2)-\frac{2\nu_a}{\nu}\frac{(1-\epsilon_0)}{\epsilon_0}. \tag{56}$$

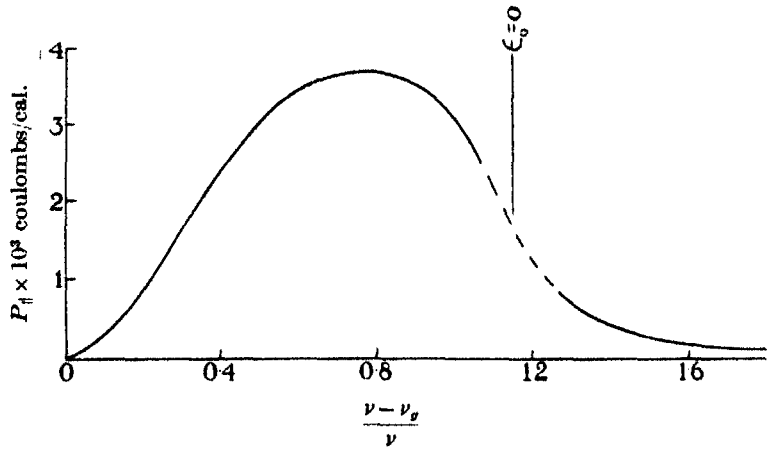


FIG. 3—Calculated spectral distribution curve for potassium. The photoelectric emission in coulombs per calorie for light incident at 50° and polarized in the plane of incidence is shown.

If, on the other hand, as Mitchell has done, we neglect div **a** (which we have seen in § 3·2 to be zero except in the now vanishingly thin transition region) in (37) and (48), the $L(k_1)$ so obtained is the same as (56) except that the last term does not appear. This result is wrong, for we see from (31) that div **a** approaches infinity throughout the transition region as this approaches zero width, and the contribution to $G(k)$ does not vanish.† Mitchell's results are therefore in error, and the term neglected is important, especially towards the frequency at which ϵ_0 vanishes.

---

† The corresponding error in the "stationary" method of calculating the current is the assumption, when ϵ is discontinuous at $x = 0$ (so that A_x is discontinuous), that the solution v_k of (37) has its gradient continuous at $x = 0$. Actually, when $\Phi = 0$ we find on integrating both sides of (37) through the small range $(-\xi, \xi)$ that

$$\left[\frac{\partial v_k}{\partial x}\right]_{-\xi}^{\xi} = -\frac{2\pi ie}{hc}[A_x]_{-\xi}^{\xi}\,u_k.$$

We may make the further criticism that in his calculations of $P_{\parallel}$ for potassium, the extinction constant κ, for which we found the value (18), is taken as that found experimentally for normal incidence at $\nu = 5\cdot1 \times 10^{14}$, namely 1·5. However, κ is seen to vary rapidly with ν and θ, so that the approximation is a poor one since $P_{\parallel}$ depends markedly on κ.

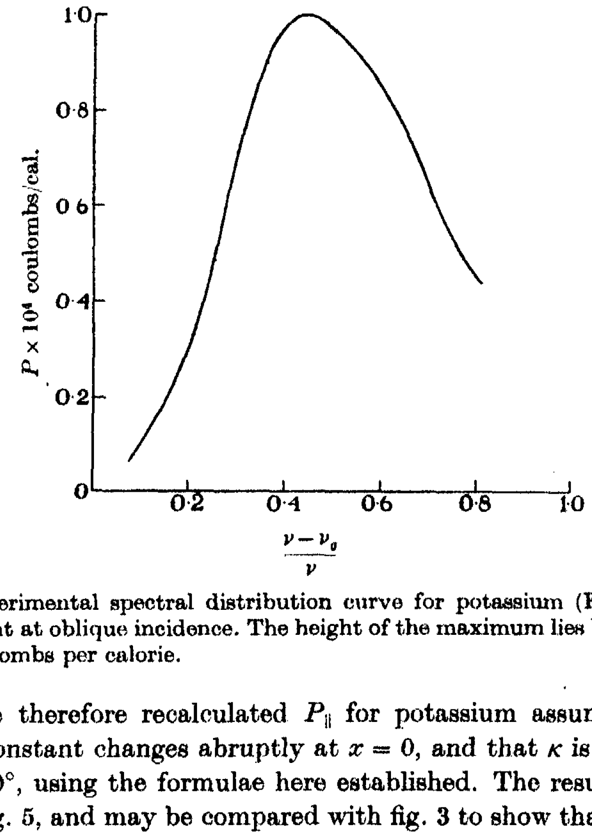


FIG. 4—Experimental spectral distribution curve for potassium (Klauer), for unpolarized light at oblique incidence. The height of the maximum lies between 0·5 and 1×10^{-4} coulombs per calorie.

We have therefore recalculated $P_{\parallel}$ for potassium assuming that the dielectric constant changes abruptly at $x = 0$, and that κ is given by (18) with $\theta = 50°$, using the formulae here established. The resulting curve is shown in fig. 5, and may be compared with fig. 3 to show that the effect of making the transition region for the optical constants of finite width is to increase the photoelectric current considerably at frequencies below ν_0. The increased values of $| L(k_1) |^2$ responsible for this arise largely because of the residue term mentioned in § 6·1.

THE EFFECT OF SLIGHT SURFACE ROUGHNESS

7—As Mitchell (1936) has pointed out, roughness of the surface will modify the above expressions for the current, in particular greatly increasing $P_{\perp}$. This is because light polarized perpendicular to the plane of incidence to the surface as a whole may yet have a component of its electric vector in the plane of incidence to an element of surface, and it is then much more effective in producing photo-emission.

We discuss briefly the effect on the currents here deduced, when the scale of the irregularities is small compared with a light wave-length but much larger than the electron wave-lengths; such irregularities will usually exist even on an optically smooth face of the metal. We suppose the surface divisible into small flat areas dS, inclined to the mean normal at the angles α, as in fig. 6, the angle β having a random distribution between 0 and 2π; this should be a fair approximation to a polycrystalline mirror surface. The electron distribution near dS will be as in fig. 1, but with x', the co-ordinate normal to dS, replacing x. We now have $\epsilon = \epsilon(x, y, z)$.

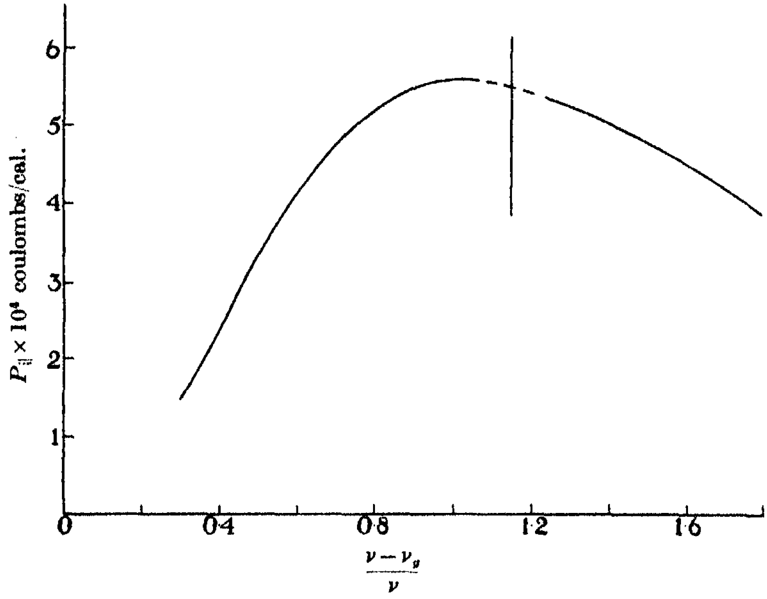


FIG. 5—Calculated spectral distribution curve for potassium, assuming that the dielectric constant changes abruptly at the surface; light incident at 50°, polarized in the plane of incidence.

7·1—Suppose that the electric vector of the incident light lies in the plane of incidence to the surface as a whole. Equations (7) now give, in place of (16),

$$\nabla^2\mathbf{H} + \frac{1}{\epsilon}\operatorname{grad}\epsilon \times \operatorname{curl}\mathbf{H} + \frac{4\pi^2\nu^2}{c^2}\epsilon\mathbf{H} = 0, \tag{57}$$

but since the variations in ϵ occupy a length much less extended in the x direction than $c/2\pi\nu$, $\mathbf{H}$ will be to a good approximation unaltered by the roughness, so we have as before $H_x = H_y = 0$, and we may put $H_z = H_z(0)$ near the surface. Since only the resolved part of the light wave polarized in the plane of incidence to dS is effective in producing emission from it, we may assume the current density normal to dS to be

$$dJ'_{\parallel}\{\theta,\, H_z(0)\} = J_{\parallel}\{\theta', H_z(0)\cos\phi\}\, dS, \tag{58}$$

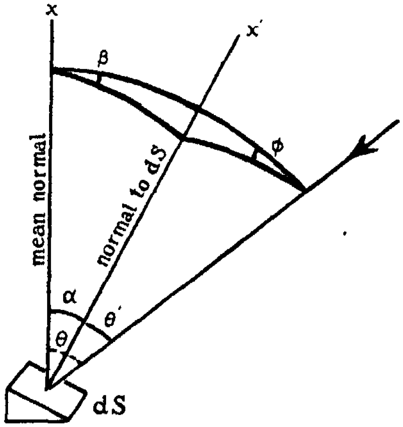


FIG. 6

where the term on the right is the current calculated from (54) for an angle of incidence θ' and with $H_z(0)\cos\phi$ replacing $H_z(0)$. In experiments the total emission of the surface is measured, so we do not take the resolved part of $dJ'_{\parallel}$ in the x direction, as Mitchell (1936) has done. We neglect emitted electrons which re-enter the surface, their number being small if the mean value of α is not too large.

If $d\sigma$ is the area of the projection of dS on the mean plane of the surface, the current from unit area of this plane is then

$$J'_{\parallel}\{\theta, H_z(0)\} = \int \frac{J_{\parallel}\{\theta' H_z(0)\cos\phi\}}{\cos\alpha}\, d\sigma,$$

the integral being taken over unit area. Now by (54)

$$J_{\parallel}\{\theta', H_z(0)\cos\phi\} = -\sin^2\theta' \cos^2\phi\, H_z(0)^2\, F(\nu),$$

where
$$F(\nu) = \frac{e^3}{64\pi^5 m^2 \nu_a \nu^2} \int_{0,\,[\mu(\nu_a-\nu)]^{\frac{1}{2}}}^{k_0} \frac{r k_1^2 (k_0^2 - k_1^2)}{(q+r)^2}\, |\, L(k_1)\,|^2\, dk_1. \tag{59}$$

Hence
$$J'_{\parallel}\{\theta, H_z(0)\} = -F(\nu)\, H_z(0)^2 \int \frac{\cos^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma. \tag{60}$$

As for a perfectly smooth surface, the energy in calories incident per second on a square cm. of the mean plane is given by (27). Thus the yield in coulombs per calorie is

$$P'_{\parallel} = \frac{32\pi \times 4{\cdot}18 \times 10^8}{c^2} \frac{\epsilon_0^2 \cos\theta}{(1-\epsilon_0)\{1-(1+\epsilon_0)\cos^2\theta\}} F(\nu) \int \frac{\cos^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma \tag{61}$$

$$= \frac{P_{\parallel}}{\sin^2\theta} \int \frac{\cos^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma. \tag{62}$$

If the mean value of α is small, (62) shows that $P'_{\parallel}$ is approximately equal to $P_{\parallel}$ (the yield from a perfectly smooth surface).

7·2—If the light is polarized with the electric vector perpendicular to the plane of incidence to the surface as a whole, it may be shown from (7) that, as for a smooth surface, E_z and H_y vary only slowly over the surface region, since the scale of the fluctuations in ϵ is much less than $c/2\pi\nu$. There will be, however, altered vector potentials in this region, given by § 3·2 on replacing $H_z(0)$ there by $H_y(0) \sin\phi$ (the component of H_y perpendicular to the plane of incidence to dS), and x, θ by x', θ'.

Thus, the current density normal to dS is

$$J'_{\perp}\{\theta, E_z(0)\} = \int \frac{J_{\parallel}\{\theta', H_y(0)\sin\phi\}}{\cos\alpha}\, d\sigma,$$

where $J_{\parallel}\{\theta', H_y(0)\sin\phi\}$ is the value of $J_{\parallel}$ calculated from (54) when the above substitution is made; the integral is taken over unit area of the mean plane. Thus

$$J'_{\perp}\{\theta, E_z(0)\} = -F(\nu)\, H_y(0)^2 \int \frac{\sin^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma. \tag{63}$$

The rate of incidence of energy is given, as for a smooth surface, by (35), and since

$$\kappa E_z(0) = H_y(0),$$

the yield in coulombs per calorie is

$$P'_{\perp} = \frac{32\pi \times 4{\cdot}18 \times 10^8}{c^2} \frac{\kappa^2 \cos\theta}{1-\epsilon_0} F(\nu) \int \frac{\sin^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma. \tag{64}$$

Hence

$$\frac{P'_{\parallel}}{P'_{\perp}} = \frac{\epsilon_0^2}{\kappa^2\{1-(1+\epsilon_0)\cos^2\theta\}} \int \frac{\cos^2\phi \sin^2\theta'}{\cos\alpha} \bigg/ \int \frac{\sin^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma, \tag{65}$$

and

$$\frac{P'_{\perp}}{P_{\parallel}} = \frac{\kappa^2\{1-(1+\epsilon_0)\cos^2\theta\}}{\epsilon_0^2 \sin^2\theta} \int \frac{\sin^2\phi \sin^2\theta'}{\cos\alpha}\, d\sigma. \tag{66}$$

7·3—Since roughness of the kind here treated seems inevitable in experimental surfaces, we need to have some knowledge of the factors

$$R_{\parallel} = \int \frac{\cos^2\phi \sin^2\theta'}{\cos\alpha} d\sigma$$

and

$$R_{\parallel} = \int \frac{\sin^2\phi \sin^2\theta'}{\cos\alpha} d\sigma.$$

These will vary with the particular surface, but we may assume that $\bar{\alpha}$, the mean value of α, is of the order 15°. Then, if θ is larger than about 50°, $R_{\parallel}$ reduces to approximately $\sin^2\theta$, and $P'_{\parallel} \simeq P_{\parallel}$, while $R_{\perp}$ becomes

$$\frac{\sin^2\bar{\alpha}}{2\cos\bar{\alpha}}. \tag{67}$$

When $\theta = 0$, $R_{\parallel}$ and $R_{\perp}$ are both given approximately by (67); we have then for the yield, provided that $\epsilon_0 < 0$ so that the wave system is stationary, and if $\bar{\alpha}$ is not too large,

$$P'_{\parallel} = P'_{\perp} = \frac{32\pi \times 4{\cdot}18 \times 10^8}{c^2} \frac{(-\epsilon_0)}{1-\epsilon_0} F(\nu) \frac{\sin^2\bar{\alpha}}{2\cos\bar{\alpha}}. \tag{68}$$

In principle $\bar{\alpha}$ may be determined by fitting (64) or (65) to the experimentally found angular selectivities at any particular frequency.

The dependence on θ should be approximately as

$$\sin^2\theta\cos\theta/\{1-(1+\epsilon_0)\cos^2\theta\}$$

for $P'_{\parallel}$, and as $\cos\theta(\sin^2\theta-\epsilon_0)$ for $P'_{\perp}$. The frequency selectivity of $P'_{\perp}$ calculated from (64) is shown in fig. 7, where $P'_{\perp}/R_{\perp}$ is plotted for $\theta = 50°$.

DISCUSSION

8—It is difficult to find a basis for close comparison of the theory with reliable experiments. The most recent results for thoroughly outgassed potassium are those of Klauer (1934) with unpolarized light at an angle of incidence presumably about 50°, shown in fig. 4. He concludes that the comparatively high selective maxima found previously were due to gas contaminating the surface. The absolute magnitudes of the yields he obtained were not, however, measured precisely, being between 5 and 10×10^{-5} coulombs per calorie of incident energy at the maximum, the latter value being used in fig. 4. The calculated value of the yield P for unpolarized light will be $\frac{1}{2}(P'_{\parallel}+P'_{\perp})$, but, since the roughness constants are unknown, $P'_{\perp}$ cannot be given precisely. However, it is reasonable to compare $\frac{1}{2}P_{\parallel}$ (from fig. 3) with Klauer's curve for P. The theoretical curve is seen to be too high near the maximum by a factor of about 20, and the maximum occurs at too high a frequency.

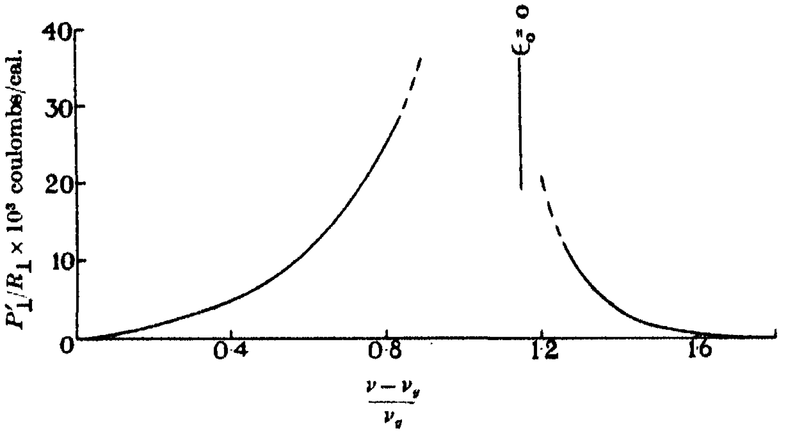


Fig. 7—Calculated spectral distribution curve for potassium, for light polarized perpendicular to the plane of incidence.

The dependence of $P'_{\parallel}$ and $P'_{\perp}$ on θ agrees with the general findings of experiment, but $P'_{\perp}$ as here calculated shows a strong selective maximum near $\nu = 10^{15}$, whereas no such effect is actually found. Equation (65) shows that $P'_{\perp}$ and $P'_{\parallel}$ have not the same dependence on frequency.

With regard to the discrepancy in the magnitude of $P_{\parallel}$, it is seen that Mitchell's suggestion that taking into account the gradual transition of the optical constants at the surface, while retaining the square potential barrier, fails to improve the agreement with experiment. In fact, as a comparison of figs. 3, 4 and 5 shows, the agreement is not as good.

There remain the following sources of error: the simple step potential barrier is a poor approximation to that actually existing, and no account has been taken of the periodicity in the yz plane of the surface field. Bardeen (1936) has found the charge distribution and the form of the barrier in the case of sodium; the barrier extends over a distance comparable with the average electron wavelength, and its effective height depends on k_1. It would not be surprising if the current given by such a model differed considerably from the present result in magnitude and frequency variation, because both the probability of excitation of an electron in the altered surface field and the probability of its escaping would be different. A similar method *of calculation is applicable, but the complication is of course increased. The yields given by fig. 3 may be expected to apply to a metal surface at which* contamination has made the barrier more like a simple step.

The experimental curves for the photoelectric yield from potassium show a rise above a frequency of about $1 \cdot 2 \times 10^{15}$. As we have seen in § 3·1, the light wave is no longer a simple decreasing exponential inside the metal when ν exceeds a limit given by (17), but will then extend for several wavelengths into the metal unless the damping is very great. We may then expect the volume photoelectric effect (Tamm and Schubin 1931) to become more important, and perhaps to be responsible for the rise mentioned above.

The author wishes to thank Mr. A. H. Wilson for suggesting this investigation, and for his constant help and interest.

SUMMARY

9—It is shown that, using the Sommerfeld model of a metal, a classical calculation of the field of an incident light wave (near the surface) gives to a close approximation the same results as the quantum theory of metallic reflexion of Schiff and Thomas. An expression is obtained for the photoelectric current arising at the surface of such a metal, taking account of the variation of the field of the light wave near the surface, on the lines of the theory given by Mitchell in which this variation is neglected. The calculated emission for potassium is compared with recent experimental results at frequencies near the selective maximum; various suggestions are made as to the lack of close agreement found. The effect of small-scale roughness on an optically smooth surface is also briefly discussed. Errors are pointed out in the theories of Mitchell and of Schiff and Thomas.

REFERENCES

Bardeen 1936 *Phys. Rev.* **49**, 653–63.
Foersterling 1928 "Lehrbuch der Optik." Leipzig: Hirzel.
Frehafer 1920 *Phys. Rev.* **15**, 110.
Klauer 1934 *Ann. Phys., Lpz.*, **5**, 20, 909–18.
Kronig and Groenewold 1934 *Physica, 's Grav.*, **1**, 4, 256.
Mitchell, K. 1934 *Proc. Roy. Soc.* A, **146**, 442–64.
—— 1936 *Proc. Roy. Soc.* A, **153**, 513–33.
Schiff, L. I. and Thomas, L. H. 1935 *Phys. Rev.* **47**, 860–9.
Tamm, Ig. and Schubin, S. 1931 *Z. Phys.* **68**, 97–113.
Wood, R. W. 1933 *Phys. Rev.* **44**, 353–60.

The Shape of the Continuous β-Spectrum of Thorium C . C′

By H. O. W. Richardson, *University, Liverpool*
and Alice Leigh-Smith, *King's College, London*

(*Communicated by J. Chadwick, F.R.S.—Received* 12 *June* 1937)

[Plates 21, 22]

Introduction

The theories of β-decay based on the neutrino hypothesis predict that a considerable proportion of the electrons emitted from a heavy nucleus will have low energies, owing to the Coulomb attraction between the electron and the nucleus. This prediction has been in apparent conflict with most experimental curves (Madgwick 1927; Scott 1935), which show the ordinate of the energy distribution falling to zero at the origin or even before it, thus even indicating a low energy end-point below which no β-rays are emitted. It is, however, probable that the experimental uncertainties in the methods which have been used are such that no definite conclusion can be drawn from them about the shape of the low energy end of the spectrum.

In these methods the source is deposited on a solid mounting and the emitted β-particles pass through a window in entering the detecting apparatus, which may be a counter, a cloud chamber or an ionization chamber. The window stops all β-particles below a certain energy, while those which pass through are reduced in energy and considerably scattered. These effects, which are well shown in curves given by Eddy (1928), produce a marked falling off in the observed number of β-particles of low energy. The use of a solid mounting for the source introduces opposite effects giving an increased number of slow electrons; for firstly, the fast electrons will eject slow secondary electrons from the solid mounting and, secondly, if the mounting is thick, a considerable reflexion of the primary electrons will occur with varying losses of energy inside the solid, so that the reflected spectrum will contain relatively more low energy rays.

The magnitude of the latter effects has been estimated (Richardson 1934) for a source of radium E on thin gold leaf and was shown to make a considerable contribution to the energy distribution so large that the number of secondary electrons at 10 kV is comparable with that of the primary electrons at the maximum of the distribution. Such low-energy rays should appear copiously in an apparatus capable of detecting them.

It is therefore desirable to employ a method in which the above uncertainties are eliminated. This can be done by using a radioactive vapour in an expansion chamber in which the tracks are traceable up to the point of origin in the gas. This makes both window and mounting unnecessary and has the further advantage that the emission of more than one ionizing particle from the same atom with energy exceeding a few thousand volts is immediately detectable, if the emissions occur within a sufficiently short interval of time.

THE METHOD OF THE RADIOACTIVE VAPOUR

Tracks from radium D starting in the gas were first obtained (Richardson and Leigh-Smith 1934) using the comparatively stable lead tetramethyl prepared by the Grignard method.

The method was then applied to the thorium active deposit, but too much confusion was caused by the α- and β-rays coming from atoms on the walls of the chamber. To avoid this a very deep cloud chamber was constructed of such dimensions that, when expanding down to atmospheric pressure, the long 8·6 cm. α-particles of thorium C′. D were unable to reach the central region of the chamber from the walls.

In order to clear away the large ionization due to the vapour source it was necessary to use three grids each containing eight 47 gauge copper wires stretched across the chamber and maintained at a potential difference of 570 V. This potential difference was reduced to zero near the end of each expansion. It was found necessary to allow an interval of three minutes at full pressure before each expansion in order that the alcohol vapour might have adequate time to diffuse uniformly over the large volume.

THE INTERCHANGE METHOD FOR THORIUM B AND C

As thorium B has a half-period of only 10·6 hr., and as the preparation of fresh lead and thorium B tetramethyl is tedious and dangerous, tests were made to see if vapour sources could be prepared by a direct interchange process by merely dipping a deposit of thorium $(B + C + C' + C'')$ into lead tetramethyl solution in ether (Leigh-Smith and Richardson 1935). It was naturally expected that any interchange which took place would be between the atoms of thorium B, an isotope of lead, and the lead atoms of the tetramethyl, giving a vapour containing thorium B. It was found, however, that the radioactive vapour which was obtained contained a great majority of thorium C atoms, showing that the interchange took place, initially at least, more rapidly between the atoms of thorium C, an isotope of bismuth, and the lead atoms of the tetramethyl. The experiments described in this paper are thus concerned with the disintegration of thorium C. Only two certain cases of thorium B disintegrations were detected in photographs using interchange sources.*

Experimental Details

The source of thorium (B + C + C′ + C″) was obtained by deposition on gold by recoil in an electric field in the presence of thoron. The piece of gold was then dipped in a dilute solution of $Pb(CH_3)_4$ in ether, and kept screened from the expansion chamber to reduce γ-ray effects.

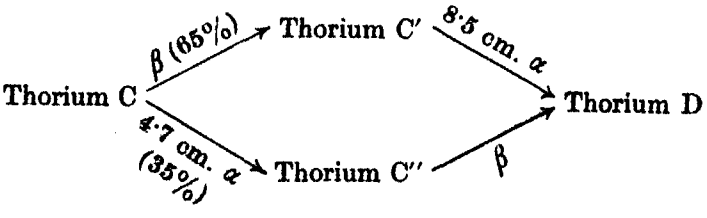
The chamber was first cleared of old droplets and oxygen-free nitrogen filtered through cotton wool was then allowed to flow in slowly. A glass rod with a cup-shaped hollow at its end was dipped into the solution and then was quickly introduced into the centre of the chamber through a hole opened in the roof. After two minutes the solution had evaporated from the rod, which was withdrawn, leaving the vapour in the centre of the chamber.

After about two slow clearing expansions the chamber was pumped up to the appropriate excess pressure and after three minutes a fast expansion was taken and photographed by means of a stereoscopic camera. Generally a new supply of thorium C vapour was introduced before alternate fast expansions.

The chamber was filled with nitrogen with a little carbon dioxide with a view to reducing the rate of oxidation of the vapour.

The Association of α-Rays with Disintegration Electrons

Owing to the branching of the thorium series at thorium C, shown in the diagram, the α-tracks starting in the gas were of two kinds. First the



---

* The thorium B disintegrations were distinguished by the fact that a pair of β-rays starting at the same point was observed, consisting of a nuclear β-ray with a secondary electron due to the internal conversion of the 238 kV γ-ray.

thorium C . C″ α-tracks of mean range 4·7 cm. and second the thorium C′D tracks of mean range 8·5 cm. Owing to the short life of $\sim 10^{-9}$ sec. of the thorium C′ nucleus the latter track always accompanied the β-track due to thorium C . C′ and served as a valuable indicator of its point of origin. Such pairs were always observed when the quality of the photograph was adequate to show the β-track. These pairs were noticed by Feather (1929) from thorium C on gold leaf. About 420 expansions were photographed in which 200 thorium C . C′ disintegration electrons were observed.

Owing to the comparatively long half-period of thorium C″ . D, viz. 3·1 min., no association between the thorium C . C″ α-particles and the thorium C″ . D β-ray was observed, the ions due to the α-track having been removed from the expansion chamber before the β-track could be photographed, even if the thorium C″ atom had not moved appreciably from its initial position in the chamber during its life-time. It is of interest that this method appears to be the only one giving directly the β-emission of thorium C . C′ unmixed with that of thorium C″ . D. As the latter emission can be obtained in the unmixed form by collecting thorium C″ by recoil, it has been previously necessary to deduce the thorium C . C′ distribution by subtraction.

The Secondary β-Rays of Thorium C.C″

Many cases were observed in which an α-track and a slow β-track started at the same point. In all such cases the α-track was reconstructed stereoscopically and its length was measured. If the observed length exceeded 5·0 cm. in standard air, the disintegration was ascribed to thorium C . C′. Three disintegrations were allotted to thorium C . C′ in which less than 5·0 cm. was visible, the nature of the α-tracks being inferred from the number of δ-rays along them.

The shorter α-tracks present were due to the disintegration of thorium C . C″ and 108 cases were observed in which the short 4·7 α-tracks of thorium C . C″ started at the same point as a slow β-ray such as is shown in fig. 3*a*, Plate 21. These β-rays must arise from the internal conversion of the 40 kV γ-ray whose association with the thorium C″ nucleus was shown experimentally by Ellis (1932) by a recoil method. This verified the prediction of Gamow (1930) that the γ-ray arose from transitions between the two states of thorium C″ resulting from the emission of the two strongest groups in the α-ray fine structure of thorium C . C″. Additional proof of the association between the 40 kV γ-ray and the thorium C . C″ α-transition is provided by the present photographs.

A rough preliminary survey was made to estimate, if possible, the proportion of the 4·7 α-tracks which were accompanied by these secondary electrons. It was necessary to include only those α-tracks in which both ends were clearly visible so that the presence or absence of a β-ray was obvious at either end. About 33 apparently unaccompanied short α-tracks were found but the number of comparable α-tracks with secondary electrons was only about 51, the remainder of the total of 108 failing to satisfy the criterion that both ends of the track should be strongly illuminated.

The α-ray fine structure intensities indicate that the 40 kV excited state of thorium C″ is excited in about 70 % of the disintegrations. Assuming that 60·7 % of the α-rays are accompanied by secondaries, this means that 86·7 % of the 40 kV γ-transitions lead to internal conversion. This may be compared with the corresponding figure of about 97 % found for the 47 kV γ-ray of radium E. To get a similar figure for thorium C″ it would be necessary to include 19 more α-tracks with secondaries. Such a change is probably within the limits of experimental error. Another possible source of error arises from the greater difficulty of seeing the relatively weakly ionizing secondaries from the outer levels of the atom in comparison with those of less energy and therefore greater specific ionization which come from the L level with about 23 kV energy. It is possible that some of the 33 apparently unaccompanied tracks had very faint secondaries of about 40 kV energy which were missed.

The Secondary β-Rays of Thorium C . C′

In contrast with thorium C″, the excitation of the thorium C′ nucleus is small, the three excited states being estimated by Ellis (1934*a*) to give rise to only about one secondary electron in 1000 disintegrations. Tracks due to such secondary electrons would accompany the nuclear β-track from the same atom. In the present photographs there are six cases of two β-tracks accompanying a thorium C′ α-track, but in all of the six cases one of the pair of β-tracks is too short to be a secondary electron arising from one of the known γ-rays. In four of these cases this short curved track was about 1 mm. long and might therefore be a δ-ray due to a close collision between the α-particle and an electron. In two cases the length was about 2 mm. and another explanation is required, such as the emission of a γ-ray of about 10 kV in about 1 % of the disintegrations.

A few cases of short tracks about 3 mm. long were found unaccompanied by fast β-tracks. These again were too long to be δ-rays and were interpreted as nuclear β-rays. One is reproduced in fig. 4*b*, Plate 22.

Measuring the Tracks

The photographs were replaced in the camera and the life-size reprojected three-dimensional image was formed in the usual way. The energy of the β-ray was then estimated in one of three ways as indicated in Table I. First, in the case of fast tracks, the curvature in a known magnetic field of 850 gauss was found by choosing a part of the track which appeared free from nuclear deflexions and setting four pins on the image of points on the track. Shadows of the four pin-points were then thrown on a focusing screen using a distant point source of light. Shadows of circles of known radius were thrown on the same screen and fitted to the four points. If the four points fitted one of the circles, its radius was recorded and the energy of the track obtained from the value of $H\rho$. Such tracks are given in column 3.

TABLE I

Energy interval kV	1	2	3	Total	Weighted total
0–50	13·0	2·0	0·5	15·5	9·55
50–100	4·5	5·5	2·5	12·5	8·5
100–150	0·5	0	7·5	8·0	5·7
150–200	0	0	8·0	8·0	6·25
200–250	0	0	7·0	7·0	6·55
250–300	—	—	11·5	11·5	10·8
300–350	—	—	4·0	4·0	4·0
350–400	—	—	6·0	6·0	6·0
400–2200	0	0	117	117	117

The advantage of this method is that it is applicable to tracks not perpendicular to H and therefore having spiral paths. The value of ρ obtained is the radius of the cylinder on which such spirals lie.

In cases where it was found impossible to make all four points lie on a circle, it was generally found that the tracks showed small nuclear deflexions which were then located and the curvatures of the undeflected parts of the track were measured. Owing to the slowness of working of the chamber and the small yield of the method, about one α-β pair in two fast expansions, it was necessary to reject as little as possible of the data available and so a few tracks are included in the distribution in which, owing to the shortness of the undeflected length available, an uncertainty of about 20 % may exist in the measurement of the curvature. Nineteen disintegrations were rejected because the undeflected arc subtended on the screen was less than a minimum of 9 mm.

The inclusion of some rather uncertain tracks was justified on the ground that what was sought was a general view of the balance between the low-energy and high-energy parts of the distribution curve rather than an accurate determination of the shape of the latter, which can be obtained more readily by other methods.

The nuclear deflexions of electrons passing through nitrogen increase so rapidly when the energy falls below 100 kV that it is almost impossible to deduce the energy from the curvature alone. In agreement with Champion (1936) it was found more difficult to detect small deflexions which cause an apparent increase of curvature than those which cause a decrease. In consequence unless care is taken the magnetic curvature of slow tracks tends to be overestimated. The tracks whose energies were obtained from the range are given in Table I, col. 1. The ranges were found by measuring the distance between the pins placed on suitable points of the three-dimensional image. The energies of these tracks in electron volts were calculated from the ranges by the formula (Williams 1931):

$$E = 24{,}600\left(\frac{n_1}{n_2}\,.\,R\right)^{\frac{1}{1\cdot 82}},$$

where n_1/n_2 is the stopping power relative to oxygen at S.T.P., R is the range in cm. The stopping power of the gas was obtained by comparing the observed mean range of the thorium C′ . D α-particles with the known value in standard air.

The energies in an intermediate third class of track in col. 2 were obtained with sufficient accuracy to allot them to the appropriate relatively wide energy interval by inspection of the scattering, specific ionization and minimum visible length available. In this third class, the end-points of the tracks were not visible so that only a minimum range could be assigned, but owing to the low energy, the deflexions were too frequent to allow of a reliable measurement of the magnetic curvature. As the appearance of tracks between 30 and 100 kV varies rapidly with energy it is believed that no important uncertainty has been introduced.

The Nuclear Energy Distribution of Thorium C . C′

The energy distribution of all the tracks of thorium C . C′ exceeding 2 mm. in range is shown in fig. 1. The frequency of occurrence of low energy tracks such as reproduced in figs. 3 and 4, Plates 21, 22, at once contradicts the long maintained view that the ordinate of a β-ray energy spectrum falls to zero at or before the origin. In fact the histogram of fig. 1 suggests at first sight that the most probable emission energy is nearly zero.

It must be remembered that the tracks of slow electrons are more easily observed than those of fast electrons owing to their much greater specific ionization. With weak illumination or other unfavourable experimental conditions fast electrons are less likely to be recorded on the photographs or, if recorded, to be suitable for measurement.

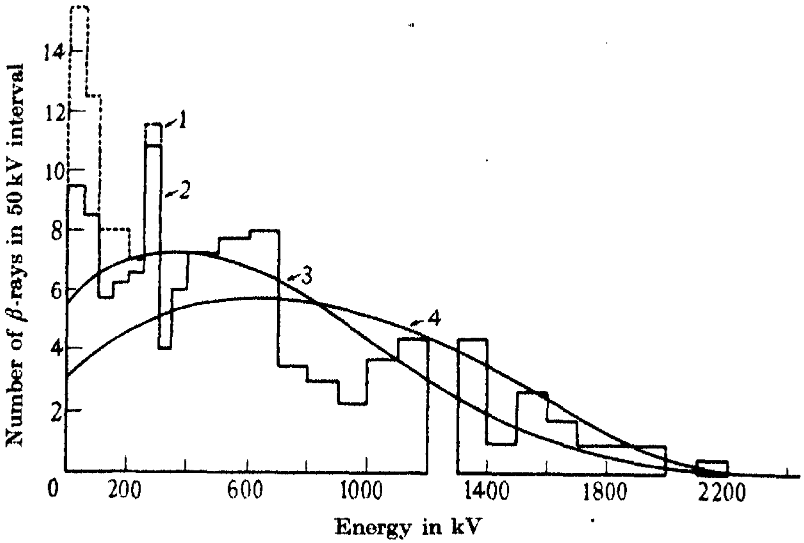


FIG. 1—Curve 1, uncorrected experimental distribution; 2, corrected experimental distribution; 3, (0, 1) distribution with $E_0 = 2600$ kV; 4, (0, 0) distribution with $E_0 = 2250$ kV.

An attempt has been made to correct for this effect by taking each slow track in turn and estimating the probability that a fast track, with its ions separated by visible intervals, would have been measurable, that is, would have subtended in the measuring screen an arc exceeding 2 cm. in length, the imaginary fast track being at the same point and in the same direction as the slow track actually seen.

In some cases it was possible to say definitely that a given track either did or did not satisfy this rather loose criterion, but in more cases it was felt safer to assign intermediate probabilities rather than certainties. It was found subsequently that the result obtained by averaging these probabilities was substantially the same as that got by giving weights of 0 and 100 % to all tracks below and above 50 %. The detailed effect of this weighting is shown in Table I. It is thought that the uncertainty arising from this general tendency to overestimate the slow tracks is smaller than from the statistical fluctuations.

Comparison with the Neutrino Theory of β-Disintegration

The presence of numerous low-energy electrons extending copiously down to the lowest detectable energy in the neighbourhood of 10 kV is in agreement with the theory which attributes their large numbers to the effect of the Coulomb attraction of the nucleus. This low energy emission is probably a general property of the β-decay of heavy nuclei. Its presence has also been established, though under less definite conditions, for radium E deposited on thin films (Richardson 1934; Alichanow, Alichanian and Dzelepow 1936). Previous failures to find the low-energy electrons are to be attributed mainly to the presence of absorbers between the source and detector.

It is of some interest to see if the results are sufficiently detailed to distinguish between the different types of interaction between heavy and light particles which have been proposed. In the original or (0, 0) formula of Fermi (1934) the probability of a given electron energy depends only upon the volume available in the momentum space of the electron and neutrino. In the (0, 1) interaction proposed by Konopinski and Uhlenbeck (1935*a*) the probability is biased in favour of large values of the neutrino momentum by introducing the first derivative of the neutrino wave function so that much more of the β-disintegration energy is, on the average, carried off by the neutrinos (or antineutrinos) than by the electrons.

To a close approximation the formulae predicted by the two hypotheses are:

$$(0,0) \quad P(E)\,dE = A(E+511)(1+0\cdot355\eta)(E_0-E)^2\,dE,$$

$$(0,1) \quad P(E)\,dE = A(E+511)(1+0\cdot355\eta)(E_0-E)^4\,dE.$$

$P(E)\,dE$ is the probability of emission of an electron with energy between E and $E+dE$ in kV, E_0 is the maximum emission energy, η is the momentum of the electron in units of $m_0c = H\rho/1700$ gauss cm.

Uhlenbeck (1936) has shown that there is but little difference between the shape of the distributions for permitted and forbidden β-transitions.

The second formula predicts more low energy electrons than the (0, 0) formula and an extremely fine tail at the high energy end of the distribution owing to the higher power of (E_0-E).

The (0, 1) formula has been shown to give a fairly good representation of the high energy parts of many β-spectra. This can be conveniently shown by the method of Kurie, J. R. Richardson and Paxton (1936) by plotting, as in fig. 2, $\left(\dfrac{N}{(1+0\cdot355\eta)}\right)^{\frac{1}{4}}$ against E where N is the number of electrons lying between η and $\eta+d\eta$, $d\eta$ in fig. 2 being 850 gauss cm.

If the (0, 1) formula is satisfied, a straight line should result cutting the energy axis at $E = E_0$.

For thorium C . C′ the points obtained lie roughly on a straight line cutting the axis of E between 2600 and 2900 kV. As the two lowest points only represent 3·5 and 1·0 tracks respectively they possess little statistical weight so that the line might well be drawn to $E_0 = 2600$ kV. The four low energy points marked by triangles represent the effect of correcting the points marked by circles for the partial β-spectra as described below, and it can be seen that the corrected points lie on the straight line down to 15 kV electron energy suggesting that, within experimental error, the formula fits down to the lowest energies.

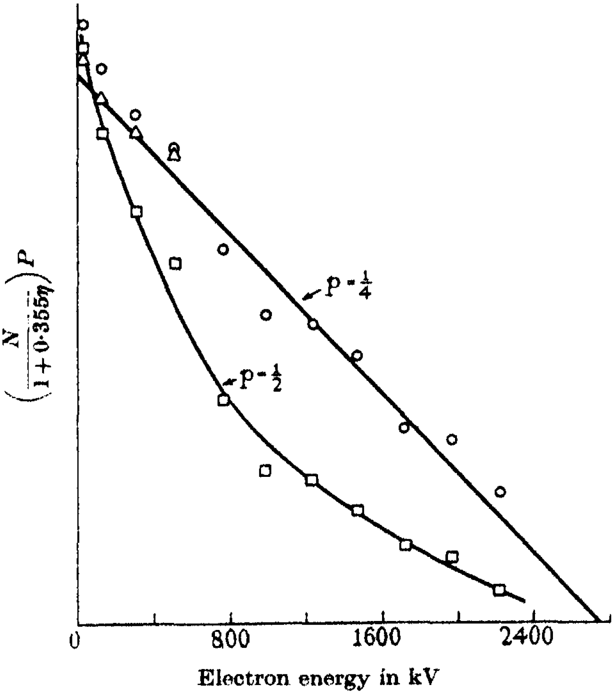


FIG. 2

If, on the other hand, the (0, 0) formula is correct, a straight line should be obtained by plotting $\left(\frac{N}{(1+0\cdot355\eta)}\right)^{\frac{1}{2}}$ against E as shown by the points marked by squares in fig. 2.

It will be seen, in agreement with results obtained for other β-disintegrations, that these points do not lie on a straight line.

The curves predicted by the (0, 0) and (0, 1) interactions with E_0 at 2250 and 2600 kV respectively are also shown in fig. 1. In order to reduce the effect of statistical fluctuations the predictions of the various distributions in wider intervals of 400 kV are compared with experiment in Tables II and III.

TABLE II—ENERGY DISTRIBUTIONS COMPARED

Energy interval	Experiment number of electrons	Theory: number of electrons			
		$E_0 = 2·25$ mV (0, 0)	$E_0 = 2·6$ mV (0, 1)	$E_0 = 2·3$ mV (0, 1)	$E_0 = 2·7$ mV (0, 1)
0–400	57·3	36·0	54·5	64·5	53·3
400–800	53·0	45·5	53·2	56·5	51·5
800–1200	27·0	42·3	35·2	31·5	34·9
1200–1600	16·5	28·2	16·4	11·1	17·6
1600–2000	9·5	11·4	4·76	1·65	5·9
2000–2400	1·0	1·0	0·5	0	1·0

TABLE III—THEORETICAL DISTRIBUTIONS WITH PARTIAL SPECTRA

Energy interval	$E_0 = 2·25$ mV (0, 0)	Deviations from Exp.	$E_0 = 2·6$ mV (0, 1)	Deviations from Exp.
0–400	43·0	− 14·3	60·4	+ 3·1
400–800	43·4	− 9·6	50·7	− 2·3
800–1200	39·8	+ 12·8	33·1	+ 5·1
1200–1600	26·5	+ 10·0	15·4	− 1·1
1600–2000	10·7	+ 1·2	4·5	− 5·0
2000–2400	0·94	− 0·1	0·5	− 0·5

The High Energy End-point and the Expression for the Interaction Energy

The "intercept" value of E_0, given by the (0, 1) line, may be taken as 2700 kV with a probable error of about ± 100 kV. This is 450 kV greater than the "experimental" value of 2250 kV obtained by Henderson (1934). Similar differences between values of E_0 obtained by the intercept method and the energies of the fastest observed tracks have been noted in other measurements of β-distributions, in particular by Lyman (1937). In the present measurements of 180 disintegrations the fastest accurately measurable electron had an energy of about 2230 kV. The results of Lyman indicate that the experimental points for $p = \frac{1}{4}$ should fit the straight line of fig. 2 until near $E = E_0$, and then follow a steeper curve reaching zero emission at a value of E a few hundred kV below the "intercept" value, E_0.

It has been pointed out (Richardson 1937) that the above result can be reproduced by a linear combination of two terms of the (0, 0) and (0, 1) form respectively, each term using the same end-point corresponding to the experimental end-point given by the fastest observed electrons. This is equivalent to assuming that the expression for the interaction energy between the heavy particles in the nucleus and the electron-neutrino field is a linear combination of two terms, the first involving the wave functions of the electron and neutrino and the second the first derivative of the neutrino wave function as well.

It is of interest to note that measurements of the low energy end of the spectrum furnish a test of the presence in the interaction energy of further terms involving higher derivatives of the wave functions. Linear combinations of such terms have been shown by Konopinski and Uhlenbeck (1935*b*) to be theoretically possible and their presence in small amounts has been suggested to explain neutron-proton attractive forces as arising from exchange of charge due to β-emission and capture.

For instance, we might assume a linear combination giving a composite distribution in which 33 % of the electrons obey a (1, 2) interaction, 53 % a (0, 1) and 14 % a (0, 0) interaction. Owing to the presence of higher derivatives of the wave functions of both neutrino and electron in the (1, 2) interaction there would be a deficiency of both high- and low-energy electrons and though the former deficiency would be quite small the latter might give an observable fall below the straight line of fig. 2 near zero energy in the case of a heavy nucleus. The fact that no deficiency of low energy electrons is apparent in the present results suggests that the coefficient of the (1, 2) interaction must be small.

The Partial β-Spectra of Thorium C . C′

The number of low-energy β-rays will be increased owing to the presence of partial β-spectra of low energy giving rise to the excited states of the thorium C′ nucleus in accordance with the theory of Ellis and Mott (1933). The effect of the partial spectra of the states of 1797 and 1623 kV excess energy which, according to Ellis (1934*b*), are excited in 2 and 4 % of the thorium C . C′ disintegrations, has been calculated assuming that the partial spectra are (0, 0) in shape with end-points at 450 and 640 kV. The results are shown in Table III and fig. 2. The corrections are small because only 6 % of the electrons are affected. The (0, 1) distribution shows the better agreement with experiment. The effect of the 726 kV excited state on the distribution is negligible.

In conclusion we wish to thank Professor C. D. Ellis for his valued encouragement and interest in the work and also Bedford College, London, for providing some radio-thorium.

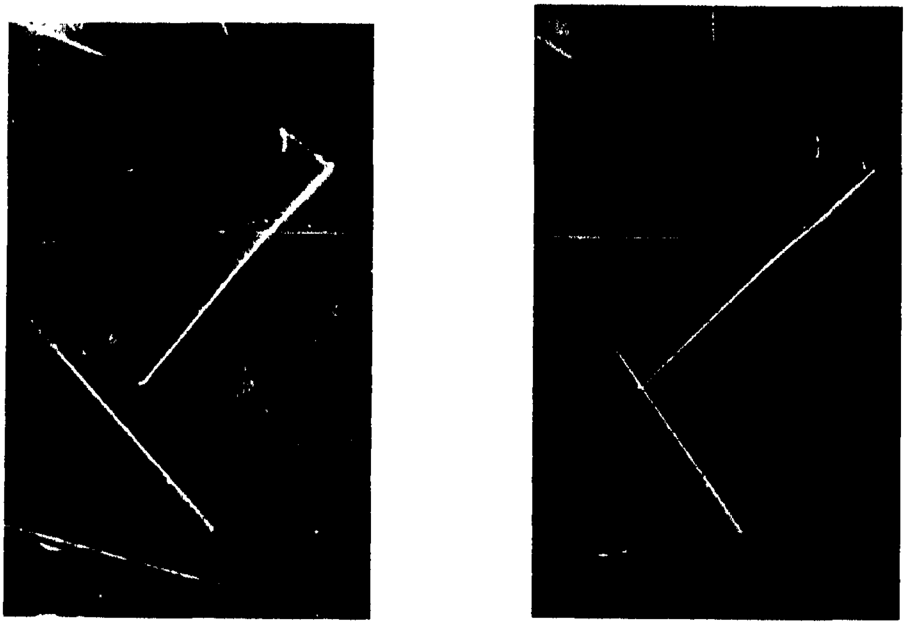


FIG. 3*a*—Thorium C.C′ α-track associated with secondary β-track due to the 40 kV γ-ray.

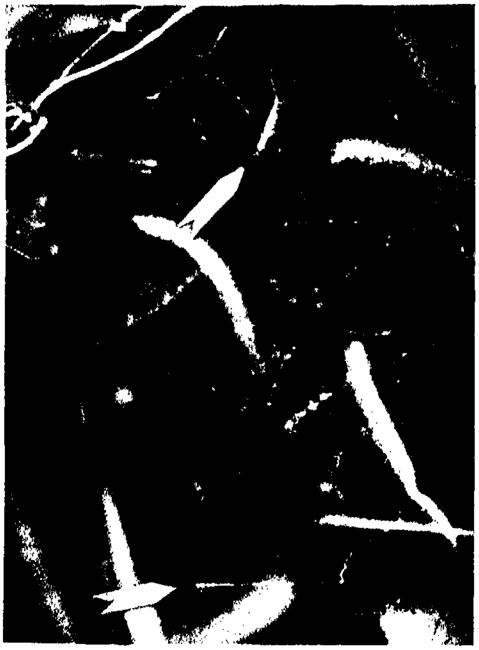


FIG. 3*b*—Two low energy nuclear β-rays of thorium C.C′.

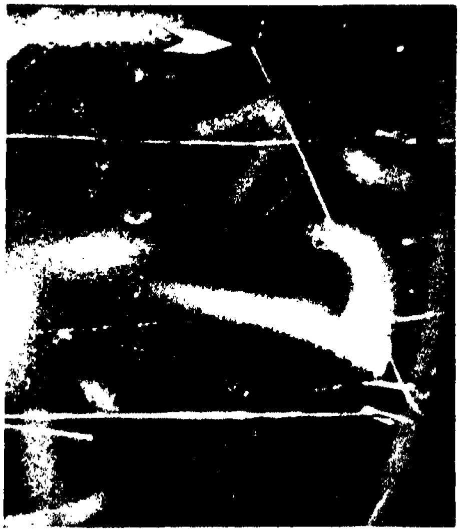


FIG. 3*c*—Nuclear β-ray of about 120 kV showing scattering.

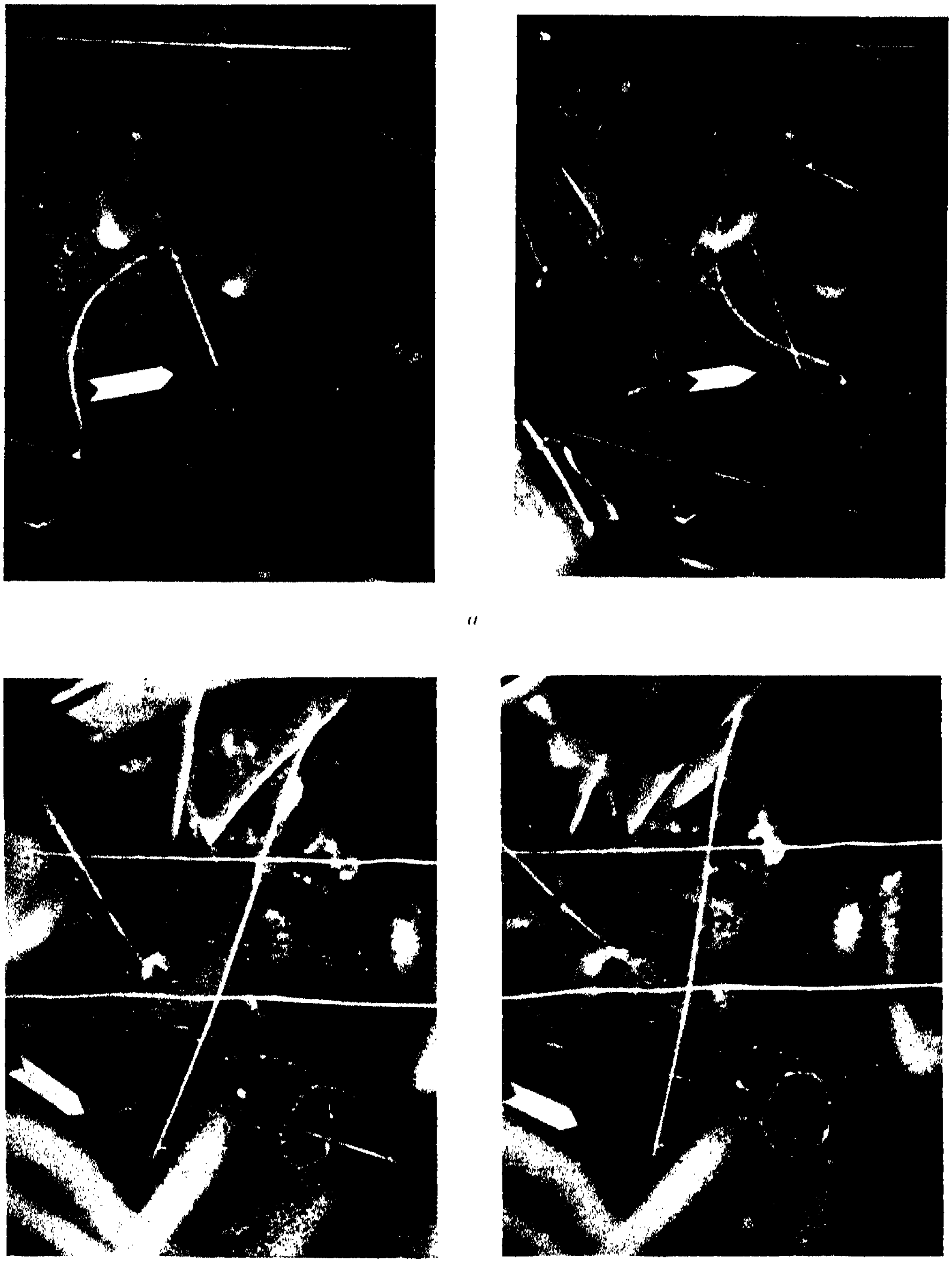


Fig. 4—Nuclear electrons of thorium C,C′ of low energy. *a*, of 40 kV; *b*, of 12 kV.

SUMMARY

A method is described by which a radioactive vapour of thorium (B + C) was introduced into a cloud chamber.

About 200 thorium C . C′ disintegrations occurring in the gas were photographed and the energies of the nuclear electrons were estimated. A large proportion of low energy electrons was found extending down to the lowest detectable energy of about 15 kV. It is concluded that the energy distribution curve does not pass through the origin but passes well above it, showing no marked fall near zero energy. The curve agrees within experimental error down to zero energy with the shape predicted by the (0, 1) interaction formula of Konopinski and Uhlenbeck (1935). The maximum electron energy obtained by fitting the data to the (0, 1) formula is 2700 ± 100 kV.

The possible presence of higher order derivatives in the interaction energy is discussed.

Numerous secondary electrons were observed associated with the α-rays of thorium C . C″.

REFERENCES

Alichanow, A. I., Alichanian, A. I. and Dzelepow, B. S. 1936 *Nature, Lond.*, **137**, 314–15.
Champion, F. C. 1936 *Proc. Roy. Soc.* A, **153**, 353–8.
Champion, F. C. and Alexander, N. S. 1936 *Nature, Lond.*, **137**, 744.
Eddy, C. E. 1928 *Proc. Camb. Phil. Soc.* **25**, 50–61.
Ellis, C. D. 1932 *Proc. Roy. Soc.* A, **136**, 396–406.
— 1934*a* *Proc. Roy. Soc.* A, **143**, 350–7.
— 1934*b* *Int. Conf. on Physics*, 1, 43–59. The Physical Soc.
Ellis, C. D. and Mott, N. F. 1933 *Proc. Roy. Soc.* A, **141**, 502–11.
Feather, N. 1929 *Proc. Camb. Phil. Soc.* **25**, 522–9.
Fermi, E. 1934 *Z. Phys.* **88**, 161–77.
Gamow, G. 1930 *Nature, Lond.*, **126**, 396.
Henderson, W. J. 1934 *Proc. Roy. Soc.* A, **147**, 572–82.
Konopinski, E. J. and Uhlenbeck, G. E. 1935*a* *Phys. Rev.* **48**, 7–12.
— — 1935*b* *Phys. Rev.* **48**, 107–8.
Kurie, F. N. D., Richardson, J. R. and Paxton, H. C. 1936 *Phys. Rev.* **49**, 368–81.
Leigh-Smith, A. and Richardson, H. O. W. 1935 *Nature, Lond.*, **135**, 828–9.
Lyman, E. M. 1937 *Phys. Rev.* **51**, 1–7.
Madgwick, E. 1927 *Proc. Camb. Phil. Soc.* **23**, 982–4.
Richardson, H. O. W. 1934 *Proc. Roy. Soc.* A, **147**, 442–54.
— 1937 *Nature, Lond.*, **139**, 505.
Richardson, H. O. W. and Leigh-Smith, A. 1934 *Nature, Lond.*, **134**, 772.
Scott, F. A. 1935 *Phys. Rev.* **48**, 391–5.
Uhlenbeck, G. E. 1936 July Colloquium, Kopenhagen (unpublished).
Williams, E. J. 1931 *Proc. Roy. Soc.* A, **130**, 310–27.

Experiments on Iron Ammonium Alum in Alternating Fields at very Low Temperatures

By A. H. Cooke and R. A. Hull

(*Communicated by F. A. Lindemann, F.R.S.—Received* 18 *June* 1937)

It has been pointed out by Simon (1935*a*) that calorimetry in the region below 1° K. can be made sensitive enough to measure accurately the thermal energy changes due to the emission and absorption of radioactive radiations. With a view to carrying out investigations of this kind, experiments have been in progress in the Clarendon Laboratory to improve the sensitivity of the method of temperature measurement. Temperatures below 1° K. are determined primarily by extrapolating Curie's law, verified at higher temperatures, which is that the susceptibility of certain paramagnetic substances is inversely proportional to their temperature. Generally the thermometric substance is the actual paramagnetic salt used to obtain the low temperature.

There are various methods of measuring the susceptibility; for instance, the substance may be surrounded by two coils whose mutual inductance is measured at the various temperatures. For this purpose A.C. methods of measuring inductance have certain advantages.* In particular a null method may be used by balancing with an external adjustable negative mutual inductance, and headphones can be used as the detecting instrument if the frequency lies in the audible ranges. There is also the possibility of compensating the effect of the unavoidable heating of the substance, the normal drift, by some form of automatically increasing mutual inductance.

It has been found, however (Giauque and MacDougall 1935*b*; MacDougall and Giauque 1936; and Gorter 1936*a*, *b*) that at low temperatures an appreciable heating is produced in paramagnetic substances by the action of the A.C. The heating must be due to an irreversible exchange of energy between the alternating field and the magnetic ions, arising from a lag between the instantaneous positions of the vectors corresponding to the field and the magnetic moment of the ions. If the A.C. method of measuring temperature is to be used in calorimetric experiments, it is necessary to eliminate this irreversible heating if possible, or, more precisely, to reduce it to such an extent that it is negligible compared with the normal drift. In order therefore to ascertain how it depends on temperature and the applied frequency and amplitude, the experiments described below have been made as direct determinations of the irreversible heating in iron ammonium alum. No attempt has been made in them to use the A.C. method of temperature measurement, as obviously this would introduce an error due to the effect under investigation; instead, the inductances and temperatures were determined by the ordinary ballistic galvanometer arrangement already available.

---

* Such a method has been used by Simon in a demonstration experiment (Simon 1935*b*). An A.C. method of determining self-inductance and hence susceptibility has been described by Giauque and MacDougall (1935*a*).

Production of the Low Temperatures

The apparatus and procedure for cooling the substance are essentially those described by Kürti and Simon (1935). In some of the preliminary experiments a metal calorimeter was used, but most of them were made using a glass calorimeter (ordinary soft glass) to ensure that there was no additional heating by conduction through gas liberated from the metal walls by induced eddy currents. The general arrangement of the part of the apparatus lying below the helium liquefier is shown in fig. 1. Powdered iron alum is compressed into a cylindrical pellet *P* 3 cm. long, 0·7 cm. in diameter, which rests on a glass bead *B*. This is supported by the cap at the bottom of the calorimeter, which is a ground joint greased in the ordinary way. Helium gas can be admitted to the space immediately round the salt to produce thermal contact with the container for the thermometer calibration and during magnetization, but this space is otherwise continuously evacuated through the tube *L*. The space between the inner and outer walls of the calorimeter is filled through the tube *R* with liquid helium *H* which thus almost entirely surrounds the salt. After calibration of the thermometer in the range 4–1° K. this liquid helium is kept at 1° K. by rapid pumping. These two glass tubes *L* and *R* leading up from the calorimeter are joined by copper-to-glass seals to copper tubes coming from the helium liquefier.

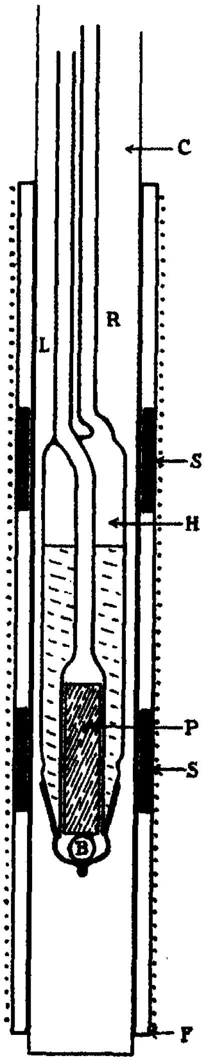


Fig. 1—The experimental arrangement.

The calorimeter is separated from the liquid hydrogen by a vacuum case

C 0·5 mm. thick made in German silver to reduce magnetic screening, and the measuring coils are wound on a bakelite former F which slips on to the vacuum case. With this arrangement the coils are permanently immersed in liquid hydrogen during an experiment so that the resistance to inductance ratio is very low. There are two identical secondaries S of 5000 turns of 47 gauge copper wire, one surrounding the salt, and one well away from it to act as compensator. Twenty-three turns per cm. of 34 gauge copper wire are wound uniformly over the whole former to act as primary. Thermometer calibration and demagnetization procedure are described in the paper (Kürti and Simon 1935) previously referred to. Large magnet pole-pieces are used to ensure that the salt is cooled uniformly.

ALTERNATING CURRENT HEATINGS

Alternating current from an oscillator and amplifier can be passed through the primary coil, which for this purpose is switched out of the ballistic circuit. A heating experiment is carried out in the following manner. The temperature drift is measured for some minutes, and then the primary circuit is switched over to the alternating current for a period of 1–5 min. The drift is afterwards measured again. Drift measurements and heatings are alternated throughout the course of the experiment.

Two frequencies were used, 455 and 328 c./sec., which were checked by comparison with tuning forks. The A.C. passed through the coil was measured by a thermal converter and millivoltmeter, and gave a field of 2–10 gauss at the pellet.

HEAT CAPACITY OF THE PELLET

To convert a measured temperature rise into heat energy it is necessary to know how the specific heat of the salt varies with temperature. This was determined in a separate experiment in which the cooled salt was heated by γ-rays; γ-ray heatings and drift measurements were alternated in a manner similar to that just described. The results obtained in this way fix the relative values of the specific heat, and can be converted into absolute values from the known specific heat of iron alum at 0·1° K. (Kürti, Lainé and Simon 1937).

THE TEMPERATURE SCALE

The temperature scale obtained by extrapolating Curie's law is to some extent arbitrary in that it depends on the material and shape of the paramagnetic sample. But the specific heats and the A.C. heatings have both been determined in terms of this same arbitrary scale and there is therefore no error from this cause in the values of the heat energy produced by the alternating current A curve showing A.C. heatings gives true values of the heat energy but in terms of arbitrary temperatures as abscissae.

Kürti, Lainé and Simon (1937) have recently extended their determination of the thermodynamic scale of temperature to temperatures well below the range used in these experiments. Their results have made it possible from knowledge of the demagnetizing factor of the pellet, calculated from its length-to-diameter ratio, to correct the arbitrary temperatures obtained here to the corresponding absolute temperatures, and in the remainder of this paper temperatures are referred to the absolute scale.

SCREENING EFFECT OF THE VACUUM CASE

In calculating the A.C. heatings in terms of the amplitude of the alternating magnetic field, it is necessary to verify that the latter is not reduced by the screening effect of the vacuum case below the value calculated from the current in the primary coil. The screening has therefore been investigated directly in a subsidiary experiment but under the conditions of the other experiments. A coil was wound to the same dimensions as the pellet and substituted for it. The mutual inductance of this coil and the primary, with and without the cold vacuum case separating them, was measured by an alternating current method using the same 455 cycle input to the primary. No effect due to the vacuum case could be detected; the diminution in amplitude produced by screening is certainly less than 1 %.

THE RESULTS

1—*Variation with Temperature*

The A.C. heating is found to rise as the temperature is lowered, the increase being more and more rapid the lower the temperature. The range used was from 0·035° K., the limit set by demagnetization with the magnet at our disposal, to 0·11° K., above which the heating was less than the normal drift and therefore too small to be measured in this way. In a typical experiment with a field of 5·9 gauss the irreversible heating produced at 0·035° was at the rate of 0·0028°/min.; at 0·11° it was 0·00015°/min.

2—*Variation with Frequency*

The heatings at the two frequencies 328 and 455 are shown in fig. 2. The ratio of these frequencies is 1·39, the square of which is 1·92. Within the experimental error the ordinates of the two curves in fig. 2 are in this ratio, which indicates that the heating is proportional to the square of the frequency.

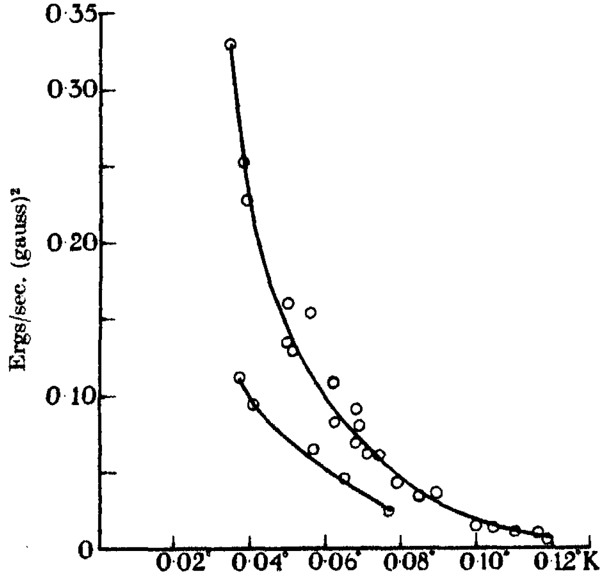


Fig. 2—Variation of the heating with temperature.

3—*Variation with Amplitude*

It is found that the heating produced varies as the square of the applied amplitude. The presence of any magnetic saturation effect would cause a decrease below these values at higher amplitudes. On the other hand, the effect of any harmonics of higher frequencies in the oscillator output would be to cause an increase at higher amplitude, as these harmonics would then be excited in greater proportion and would produce a bigger heating. But since the measured heating remains exactly proportional to the square of the amplitude for both frequencies and over the whole temperature range, it seems that this is true of the actual alternating current heating, and that saturation and harmonics are both absent.

Discussion

These results may be discussed from two points of view; as to their theoretical significance and interpretation, and in their bearing on A.C. methods of temperature measurements. Considering the former it should first be stated that an explanation of the phenomena as arising simply from magnetic hysteresis is not tenable, since this would produce a heating proportional to the first power of the frequency. It does, however, seem possible, as will now be shown, to derive a consistent explanation in terms of a time lag of the form of a relaxation time.

The Relaxation Times

Any heating effect by an alternating field can be represented to a first approximation by a lag in the phase of the magnetization. If $H = H_0 \sin \omega t$ and $\sigma = A \sin(\omega t + \phi)$, where H is the field and σ the induced magnetic moment, the irreversible heating per second is $Q = \frac{1}{2} H_0 \omega A \sin\phi$.

If the lag can be represented by a time constant of the form of a relaxation time, then $\frac{d\sigma}{dt} = \frac{1}{\tau}(\sigma_0 - \sigma)$, where $\sigma_0 = \chi H$, χ being the ordinary susceptibility. Substituting the values of H and σ previously taken, it follows that $\tan\phi = -\omega\tau$ and $A = \chi H_0 \cos\phi$. The heating per second is thus

$$\tfrac{1}{2}\chi H_0^2 \omega \sin\phi \cos\phi = \tfrac{1}{2}\chi H_0^2 \frac{\omega^2 \tau}{1+\omega^2\tau^2},$$

this expression and its derivation being analogous to that given by Debye (1934) for high frequency losses in dielectrics.* If $\omega\tau \ll 1$ this expression takes the form $\frac{1}{2}\chi H_0^2 \omega^2 \tau$.

The experimental results are of this form, and indeed they can be interpreted on this basis. For if $\frac{1}{2}\chi H_0^2 \omega^2 \tau$ is equated to the measured heating, values between 10^{-7} and 10^{-8} sec. are obtained for τ which are consistent with the assumption $\omega\tau \ll 1$. The values of τ at different temperatures are shown in fig. 3. The absolute values of the susceptibility and specific heat used in calculating them have not been determined directly for the pellet used, but have been assumed from previous determinations on similar pellets. There may be a constant error from this cause, but it will certainly not be greater than 10 %.

So far τ occurs simply as a time constant whose existence is deduced from the form of the experimental results. To find its physical origin it is necessary to know what time relations are involved in the application of a magnetic field to a paramagnetic substance. This has been considered generally by Heitler and Teller (1936). Under the conditions of the present experiments there are two time effects: the time for the magnetic ions themselves, the spins, to take up their stable orientations in the field, and the time for the energy of this magnetization to be communicated to the lattice.

---

* Compare also the derivation by Gorter and Kronig (1936, p. 1015).

The former process gives directly a lag of magnetization with respect to the applied field such as was assumed to explain the irreversible heating. The appropriate time constant t is given roughly by equating U, the energy of mutual interaction between the magnetic ions, to h/t, from which a value 10^{-9} sec. is obtained for t. This is of the same order as the value of τ found experimentally, and this process must certainly contribute to the irreversible heating.

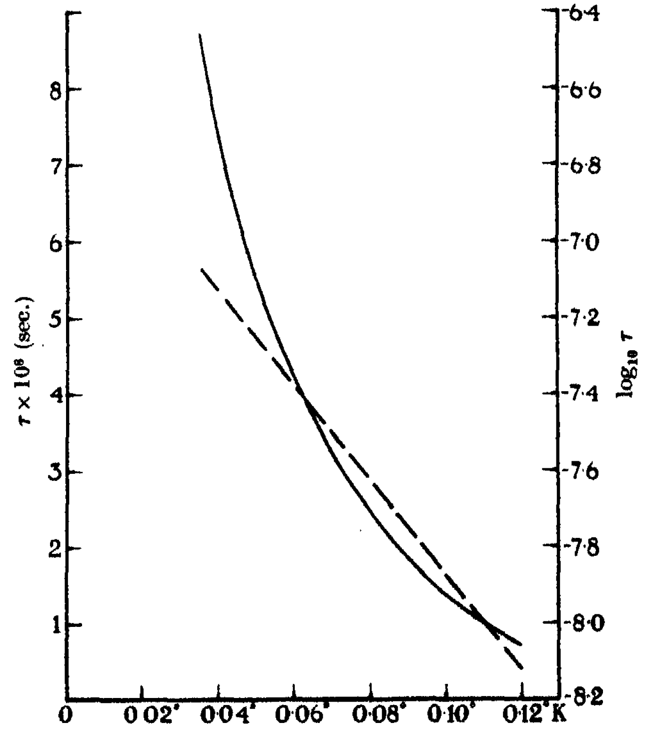


Fig. 3—Variation of the relaxation time with temperature; the ordinates on the left refer to the continuous curve, those on the right to the broken one.

The time constant Z for the establishment of thermal equilibrium between spin system and lattice is estimated by Heitler and Teller to be of the order of 40 sec. at 0·04° K. Heitler and Teller do not claim any great accuracy for this estimate, but even if it is considerably in error ωZ is now $\gg 1$, and an irreversible heating due to this second process would be independent of frequency, in contradiction to the experimental results. Also, such an irreversible heating would no longer be due to direct lag of magnetization behind the field; the magnetization could only be affected in so far as equalization of temperature between spin system and lattice reduces the temperature of the former. Now the specific heat of the spin system at these temperatures is so much greater than that of the lattice, by a factor of the order of 10^7 at 0·1° K., that the spin system temperature would not change appreciably, and it seems therefore that irreversible heating by this indirect process is negligible in comparison with that due to the direct lag of the spins themselves. The experimental values of τ may thus be interpreted as the time constant for the mutual interaction between the spins.

In fig. 3 the broken line giving the variation of $\log\tau$ with temperature is straight, which means that τ can be represented as $2{\cdot}3\times10^{-7}e^{-T/0{\cdot}035}$ sec. T occurs in this formula in a somewhat unusual form, and as at present no theoretical interpretation can be given the formula should be regarded as empirical. It should, however, be pointed out that the characteristic temperature in this formula (0·035° K.) coincides very closely with the temperature of the Curie point, found in quite an independent way (Kürti, Lainé and Simon 1937), which is also a temperature characteristic for the mutual interaction between the magnetic dipoles. This may mean that the formula has also some physical significance.*

Note on other Work

If we interpret the results of Giauque previously mentioned on the basis of a relaxation time, the value of τ obtained for gadolinium nitrobenzene sulfonate heptahydrate at 0·25°, for instance, is found to be 5×10^{-7} sec. This is comparable with the values here determined for iron alum.

The experiments of Gorter on iron alum at much higher temperatures (in the region 15–70° K.) using frequencies of about 10^7 show a relaxation time of the order of 10^{-10} sec. independent of temperature. Extrapolation of the expression $2{\cdot}3\times10^{-7}e^{-T/0{\cdot}035}$ to 15° K. gives a value 10^{-200} for τ, negligible in comparison with 10^{-10}. On the other hand, the value 10^{-10} sec. is negligible in comparison with the values of τ below 1°. It seems therefore probable that Gorter's experiments involve a different phenomenon from that interpreted here by the relaxation time τ.

---

* On the other hand, it may be that the Stark effect of the lattice which leads to a splitting of the ground state, may have an influence too on the phenomena under consideration, and that both influences together lead by chance to the experimentally observed form of the curve.

Temperature Measurement by Alternating Current Methods

From the experimental data it is evident how the effect of the relaxation heating may be minimized when measuring temperatures by alternating current methods. The following conditions must be observed:

(1) With iron alum only the region above 0·1° K. can be used. There is an apparent consequent loss of sensitivity since the decrease $\Delta\chi$ in susceptibility for a rise ΔT in temperature is given by $\Delta\chi = \frac{\chi}{T}\Delta T$, i.e. $\Delta\chi/\Delta T$ varies as $1/T^2$. This loss is, however, exactly counterbalanced by the larger ΔT for a given heat input, since the specific heat in this region falls off, with rising temperature, according to a $1/T^2$ law.

The upper limit to the temperature range available is fixed by a different consideration. The heat insulation of the cooled salt depends to a large extent on the extremely low vapour pressure of helium at these temperatures, so that any gas still present when demagnetizing is condensed on the salt. On warming the salt no appreciable loss of insulation occurs due to desorption of this gas until about 0·4° K., which is thus the upper limit to the available temperature range for high accuracy calorimetry.

(2) A lower frequency must be used. The sensitivity, understood as the voltage induced in the secondary coil, falls off proportionally to the frequency. But since the relaxation heating falls off proportionally to the square of the frequency, there is a net gain in effective sensitivity.

(3) The amplitude must of course be reduced so far as the method of measurement will permit. The experiments described were all carried out with amplitudes more than ten times that necessary for temperature measurement.

Using a magnetic field amplitude of one gauss with a frequency of 50, the calculated irreversible heating would produce a temperature rise of 4×10^{-8} degree/min. at 0·15°, which is less than the normal drift.

It may be observed in conclusion that a metal calorimeter may be used, as the early experiments with a German silver calorimeter, while not so accurate, gave results for the heatings which were certainly only of the same order as these purely relaxation heatings.

We wish to express our indebtedness to the Royal Society for a grant in aid of apparatus, and to the Department of Scientific and Industrial Research whose assistance has made this work possible for us. We have also to thank Professor F. A. Lindemann for extending to us the facilities of his laboratory, Dr. F. Simon for valuable advice and discussion and Dr. N. Kürti and Mr. G. L. Pickard for their kind help with the experiments.

SUMMARY

The irreversible heating of iron ammonium alum by an alternating magnetic field has been measured between 0·035 and 0·11° K. The heating is found to vary as the square of the field amplitude, and as the square of the frequency. It increases rapidly as the temperature is lowered.

If these results are interpreted by a relaxation time, corresponding to a lag between the applied magnetic field and the induced magnetization, it varies from 9×10^{-8} sec. at 0·035° K. to 10^{-8} sec. at 0·11° K.

It is shown that this time refers to that necessary for the magnetic ions, the spins, to take up their stable orientations in the field, rather than that necessary for the communication of the energy of magnetization from the spin system to the lattice system.

A discussion is given of the conditions under which this irreversible heating may be minimized in calorimetric experiments with alternating current methods of temperature measurement.

REFERENCES

Debye, P. 1934 *Phys. Z.* **35**, 101–6.
Giauque, W. F. and MacDougall, D. P. 1935*a* *J. Amer. Chem. Soc.* **57**, 1175–85.
— 1935*b* *Phys. Rev.* **47**, 885–6.
Gorter, C. J. 1936*a* *Nature, Lond.*, **137**, 190.
— 1936*b* *Physica, 's Grav.*, **3**, 503–14.
Gorter, C. J. and Kronig, R. de L. 1936 *Physica, 's Grav.*, **3**, 1009–20.
Heitler, W. and Teller, E. 1936 *Proc. Roy. Soc.* A, **155**, 629–39.
Kürti, N., Lainé, P. and Simon, F. 1937 *C.R. Acad. Sci., Paris*, **204**, 754–6.
Kürti, N. and Simon, F. 1935 *Proc. Roy. Soc.* A, **149**, 152–76.
MacDougall, D. P. and Giauque, W. F. 1936 *J. Amer. Chem. Soc.* **58**, 1032–7.
Simon, F. 1935*a* *Nature, Lond.*, **135**, 763.
— 1935*b* *Not. Proc. Roy. Instn.*, **28**, 515–41.

The Absorption Spectrum of Liquid Bromine

By D. Porret

From the Sir William Ramsay Laboratories of Inorganic and Physical Chemistry, University College, London, W.C. 1

(*Communicated by F. G. Donnan, F.R.S.—Received* 18 *June* 1937)

The continuous absorption spectra of gaseous bromine (Peskow 1917; Ribaud 1919; Gray and Style 1929; Acton, Aikin and Bayliss 1936) and of dissolved bromine (Bovis 1929; Gillam and Morton 1929) have been studied many times. They present a wide continuum (from about 30,000 to 17,000 cm.$^{-1}$) with a maximum at 24,000 cm.$^{-1}$. For the gas the continuum is preceded by two band systems on the long wave-length side. These systems converge at 19,585 and 15,896 cm.$^{-1}$ respectively. Acton, Aikin and Bayliss (1936) have shown that the continuum is not simple, and Mulliken (1936) and Darbyshire (1937) have pointed out that there are three overlapping continua corresponding to transitions from the ground state to three different excited electronic states. They are ${}^3\Pi_0^+ \leftarrow {}^1\Sigma_g$, ${}^3\Pi_1 \leftarrow {}^1\Sigma_g$ and ${}^1\Pi \leftarrow {}^1\Sigma_g$.

The absorption spectrum of liquid bromine has been studied by Bovis (1929) from 18,525 to 31,750 cm.$^{-1}$ and by Camichel (1893) for two frequencies only (16,978 and 18,691 cm.$^{-1}$).

The very high absorption coefficient of liquid bromine in the region of the maximum (24,000 cm.$^{-1}$) makes its measurement difficult. Because there are only a few data available on this subject, and because there is considerable interest connected with the comparison of the spectra of gaseous and liquid bromine, we considered it useful to measure again the absorption spectrum of liquid bromine over a wider range of frequencies.

Experimental

Measurements with 10, 1, and 0·1 cm. cells were made with a Judd Lewis divided beam photometer and a Hilger glass spectrograph. The light source was a 500 c.p. pointolite lamp, and Ilford rapid process panchromatic and Eastman Kodak 3L infra-red plates were used. Pairs of spectra corresponding to known optical densities were obtained and the match point determined visually. As indicated in fig. 1, this method gave points from about 12,300 to 15,200 cm.$^{-1}$.

In order to obtain points at shorter wave-lengths, very thin layers of bromine had to be used. For this purpose a wedge-shaped cell was first used.

The angle corresponds to an increase of thickness of 0·00953 mm./mm. length (Goodeve and Richardson 1937). Two horizontal lines, parallel to the bottom of the wedge, were etched on the face of the cell. The cell was placed in front of the slit of a Hilger E3 quartz spectrograph. The length of the slit was adjusted so that the reference lines and the bottom of the cell appeared in the spectrum. The light source was a hydrogen discharge tube, and Ilford ortho plates were used. Photographs were first taken with the full cell, and afterwards with the empty cell and decreasing times. The blackening of the plates was compared by means of a Zeiss microphotometer. The continuous range of thicknesses between 0·00953 and 0·14 mm. gave points from 17,000 to 18,000 cm.$^{-1}$ and further in the ultra-violet from 28,100 to 29,500 cm.$^{-1}$.

In order to get the maximum, thinner layers were obtained by pressing a drop of liquid between two optically worked quartz plates and surrounding the edges with paraffin wax. The photographs were taken with an iron tungsten spark, a divided beam Spekker photometer, and a Hilger medium quartz E3 spectrograph. As in the first system, pairs of spectra, corresponding to a known optical density, were obtained and the match point determined visually. The thickness of the bromine layer was calculated from the optical density required to get match points at wave-lengths for which the extinction coefficient was known. Three different layers of bromine were used, each of which gave match points from the maximum down to lower densities, overlapping on both sides with the wedge cell measurements.

Previous authors (Bovis 1929) prepared such layers of liquid, but they determined the thickness by means of Newton's rings. This method is not very accurate, as the thickness is of the order of one or two wave-lengths and varies from place to place. In our experiments, the light beam always crossed the bromine layer in the same spot, in the calibration as well as in the extinction coefficient measurement itself. Therefore the accuracy of the absolute values obtained depends on the reliability of the wedge cell measurements. In these, the thickness can be determined accurately, and the error for ϵ does not exceed 2 %.

RESULTS

The molecular extinction coefficient is determined by the equation

$$\epsilon = \log_{10}\frac{I_0}{I}\cdot\frac{1}{c.l},$$

where I_0 and I are the intensity of the incident and transmitted beams respectively, c the concentration in g.mol./litre, and l the thickness of the absorption layer in centimetres.

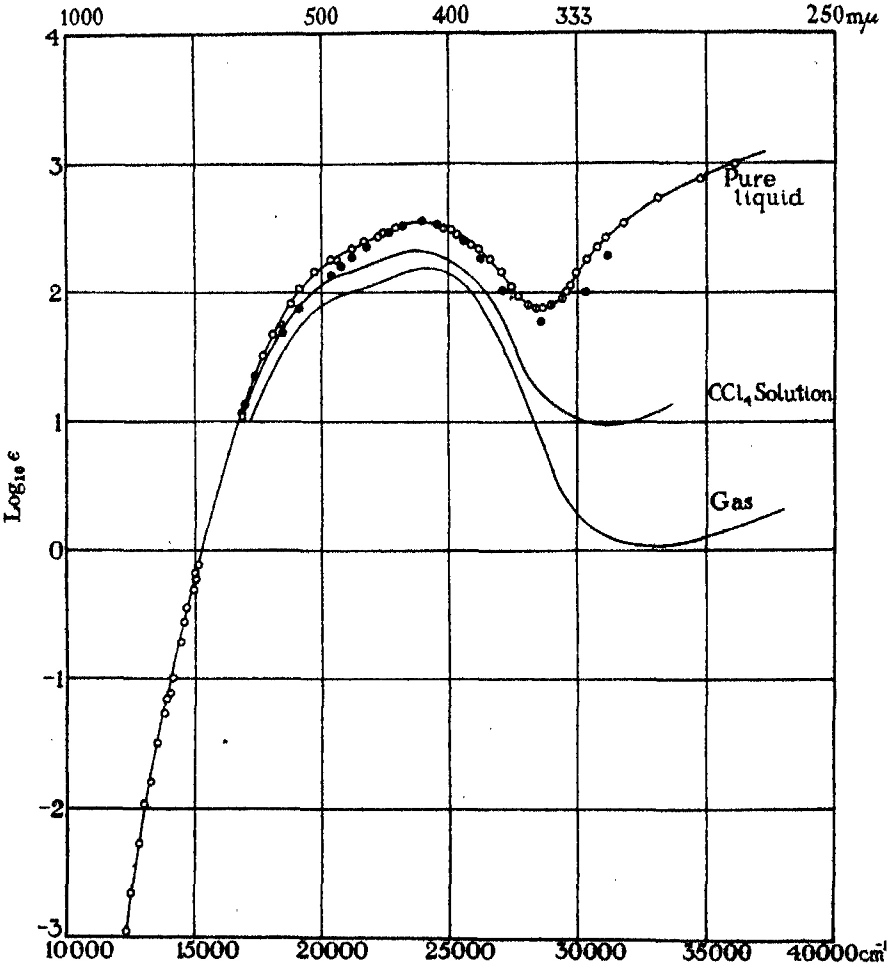


Fig. 1—□ Camichel; ● Bovis; ⦶ wedge cell.

The logarithms of ϵ are plotted against the frequency in fig. 1. The points obtained by Bovis and Camichel are also indicated. The corresponding curves of Acton, Aikin and Bayliss for the gas, and of Gillam and Morton for bromine dissolved in CCl_4, are also shown. Some results of Gray and Style were used to indicate the continuation of Acton, Aikin and Bayliss's curve beyond 30,000 cm.$^{-1}$.

DISCUSSION

It is seen from fig. 1 that the extinction coefficient curve of liquid bromine has a maximum value at 24,000 cm.$^{-1}$, followed by a minimum at 28,600 cm.$^{-1}$. It was not possible to determine the maximum of the second band because it requires too thin a layer of bromine and lies probably farther than 45,000 cm.$^{-1}$. On the long wave-length side of the maximum, the curve presents a hump at about 20,000 cm.$^{-1}$, and then goes down quite smoothly until 12,000 cm.$^{-1}$. The hump indicates that in the liquid, as well as in the gaseous and dissolved bromine, the first band is not a single one. The form of the first part of the second band also suggests that it is a composite one. The data of Bovis and of Camichel are in good agreement with ours. According to Bovis, the maximum is at 24,000 cm.$^{-1}$, but his curve does not present a hump and diverges appreciably from ours in the region of the minimum.

Fig. 1 shows that the positions of the maximum of the hump and the shape of the curve do not change appreciably for the three states considered. But in the region of the maximum, ϵ is more than twice as large for liquid as for gaseous bromine, the difference being much bigger for shorter wave-lengths. Beer's law is not valid for the passage from the gas to the dissolved and liquid states. It has been shown (Ribaud 1919) that for $\lambda < 5430$ A gaseous bromine does not follow exactly Beer's law, ϵ increasing slightly with the concentration. The same discrepancy will most probably occur with dissolved bromine. These facts show that in the three states considered absorption of light by bromine involves the same electronic transitions, but the probability of these transitions increases on passing from the gaseous to the dissolved and liquid states.

According to Mulliken (1936), the A continuum (maximum at 24,000 cm.$^{-1}$) may be attributed to the ${}^1\Pi \leftarrow {}^1\Sigma_g$ transition, and the B continuum (maximum at 20,400 cm.$^{-1}$) to the two transitions ${}^3\Pi_1 \leftarrow {}^1\Sigma_g$ and ${}^3\Pi_0^+ \leftarrow {}^1\Sigma_g$. But it is also possible to attribute A to ${}^1\Pi \leftarrow {}^1\Sigma_g$ and ${}^3\Pi_0^+ \leftarrow {}^1\Sigma_g$ and B to ${}^3\Pi_1 \leftarrow {}^1\Sigma_g$. Darbyshire (1937) supports the first interpretation of Mulliken, but he points out that one could also attribute A to ${}^3\Pi_1 \leftarrow {}^1\Sigma_g$ and B to ${}^3\Pi_0^+ \leftarrow {}^1\Sigma_g$ transitions, ${}^1\Pi \leftarrow {}^1\Sigma_g$ giving a weak continuum, the maximum lying at about 22,000 cm.$^{-1}$.

In bromine the probability of such transitions is not very high. This is shown by the relatively small absolute value of ϵ maximum (90, 160 and 740 for Cl_2, Br_2 and I_2 respectively). It is known from the Stark effect that an electric field can weaken the selection rules, i.e. increase the probability of slightly improbable transitions. It is clear that the influence of the electric fields of the surrounding molecules must increase with the concentration and that ϵ will be larger for liquid than for gaseous bromine. With the second explanation of Mulliken and the second of Darbyshire, one could more easily understand why the two maxima increase in approximately the same ratio on passing from the gas to the liquid, as in both these explanations the continua would be due mainly to triplet ← singlet transitions. Our results, however, do not allow any decisive conclusion to be drawn concerning these interpretations.

Fig. 1 shows that the minimum of the curve shifts towards the short wave lengths in passing from the liquid to the solution and to the gas, i.e. the second absorption band becomes more intense or shifts towards the red as the density of the bromine increases. The difference between ϵ for the gas and the liquid increases with the frequency and becomes very large for the far ultra-violet band. The explanation of this fact is probably that the influence of the surrounding molecules, in an electronic transition, increases with increasing excitation energy. The higher the energy of an excited state, the more strongly is the state affected by the forces between the molecules and their surroundings.

As we said before, the extinction coefficient curve falls regularly on the long wave-length side, without any sign of band structure up to the limit of the infra-red. It is a case of complete broadening of the lines. This broadening must also be due to the influence of the surrounding molecules—especially to the shortening of the lifetime of the excited molecules by so-called "induced predissociation (Kondratjew)".

In conclusion the author wishes to express his thanks to Dr. C. F. Goodeve and Dr. E. Rabinowitch for valuable discussions, and to the Swiss Ramsay Committee and the Ramsay Memorial Fellowship Trustees for the award of a Fellowship.

SUMMARY

The absorption spectrum of liquid bromine has been measured between 2,000 and 40,000 cm.$^{-1}$. In the region of the maximum, the measurements have been made by a new method, giving more accurate results than have been previously obtained. The extinction coefficient curve is compared with those of gaseous bromine and of the carbon tetrachloride solution. An interpretation of the difference between the three states is attempted.

REFERENCES

Acton, A. P., Aikin, R. G. and Bayliss, N. S. 1936 *J. Chem. Phys.* **4**, 474–9.
Bovis, P. 1929 *J. Phys. Radium* (6), **10**, 267–72.
Camichel, C. 1893 *C.R. Acad. Sci., Paris*, **117**, 307–9.
Darbyshire, O. 1937 *Proc. Roy. Soc.* A, **159**, 93–109.
Gillam, A. E. and Morton, R. A. 1929 *Proc. Roy. Soc.* A, **124**, 604–16.
Goodeve, C. F. and Richardson, F. D. 1937 *Trans. Faraday Soc.* **33**, 453–7.
Gray, L. T. M. and Style, D. W. G. 1929 *Proc. Roy. Soc.* A, **126**, 603–12.
Mulliken, R. S. 1936 *J. Chem. Phys.* **4**, 620–1.
Peskow, N. P. 1917 *J. Phys. Chem.* **21**, 382–401.
Ribaud, G. 1919 *Ann. Phys., Paris* (9), **12**, 135–226.

Investigations of Infra-red Spectra—Absorption of some Hydroxy Compounds in the Region of 3μ

BY J. J. FOX AND A. E. MARTIN

(*Communicated by Sir Robert Robertson, F.R.S.—Received* 28 *June* 1937)

INTRODUCTION

It has been realized by several investigators that with some exceptions compounds containing the hydroxyl group have a sharp absorption band close to $2 \cdot 75\mu$, and frequently another much wider band about 3μ. Errera and Mollet (1936, 1937) and Errera (1937) have shown that the latter is an "association" band which diminishes on dilution in a non-polar solvent, or by raising the temperature, when the association complexes split up and the hydroxyl band at $2 \cdot 75\mu$ becomes more pronounced. (Compare also Barchewitz (1937) and Freymann (1937).) In the present investigation of the infra-red absorption spectra of certain hydroxy compounds in the 3μ region we have made observations of the effect of the structure of the molecule on both the above-mentioned bands, as well as on the CH vibration bands.

The OH bands in the region about $1 \cdot 4$–$1 \cdot 6\mu$ have been investigated by Wulf and his co-workers (1935, 1936) for the most part with solutions of about 0·01 mol./l. in which the association band had practically disappeared. This work relates to the first overtone of the fundamental frequency in the region of 3μ, and has revealed striking constancy of wave-length and intensity of the bands for diverse compounds which may have an OH group in the molecule, and absence of such bands in certain cases where a hydrogen linkage can be postulated. Our measurements have been carried out at 18° C. on the fundamental bands about 3μ for solutions in carbon tetrachloride and other solvents at a concentration generally about 0·1 mol./l., when association is appreciable. The result of working on the fundamental bands rather than harmonics is to obtain much greater intensity of the bands, and simpler spectra generally.

EXPERIMENTAL

We employed an echelette grating ruled by Professor R. W. Wood, with 2050 lines to the inch, on a concave surface of 1 m. focal length. This was used in conjunction with the rock-salt prism spectrometer, thermopile and galvanometer system which have previously been described by Robertson, Fox and Martin (1934). The second-order spectrum was selected for the 3μ region and calibration effected by visual observation of appropriate orders of the Hg line λ 5460·7. All wave-lengths given in this work for sharp bands may be regarded as accurate to $\pm 0{\cdot}001\mu$. Slits were adjusted so that a galvanometer throw of 10–20 cm. was obtained, and under these conditions it was possible to obtain the fine structure of lines less than 2 cm.$^{-1}$ apart.

Absorption cells were constructed with fluorite plates cemented to a brass ring by means of a film of fish glue, and three sizes of cell were used varying from 1·3 to 10 mm. in length. Two similar absorption cells were selected, one containing solution and the other solvent, and these were placed alternately in the light path for measurements at each wave-length, so that reflexion and other losses would be compensated.

When the absorption is nearly complete the calculated value of the extinction coefficient is greatly influenced by the presence of a small amount of stray light in the spectrometer. This effect was found to be most pronounced on very sharp absorption bands, and it was allowed for in our calculations by observing the light apparently transmitted through a highly concentrated solution which is known to absorb completely at the desired wave-length.

All the compounds investigated and the solvents used in this work were specially purified. A.R. carbon tetrachloride was fractionated before use, a middle fraction boiling at 76·7° C./760 mm. being retained. Chloroform was repeatedly washed with distilled water to remove alcohol, dried over P_2O_5 and fractionated immediately before use.

RESULTS

OH *bands*

Phenol at various Concentrations in Carbon Tetrachloride—Fig. 1 shows the results obtained for phenol in carbon tetrachloride at four different concentrations. It will be observed that the curve for the lowest concentration (I) shows little, if any, of the association band around 2·95μ, although the other OH band at 2·769μ is still intense. This sharp OH band for the two intermediate concentrations could not be conveniently shown in fig. 1 without complicating it. For the purpose of considering the intensities of the two OH bands at different concentrations, it is convenient to work with the optical density of the solution rather than percentage absorption or transmission, the density being here defined as $d = \log_{10} I_0/I$, where I_0 is the intensity of the light passed through a cell containing solvent alone, and I the intensity of the light transmitted through the solution at a given wave-length. The molecular extinction coefficient κ is obtained from d by the relation $\kappa = d/cl$, where c is the concentration of the solution in mol./l. and l is the cell length in cm.

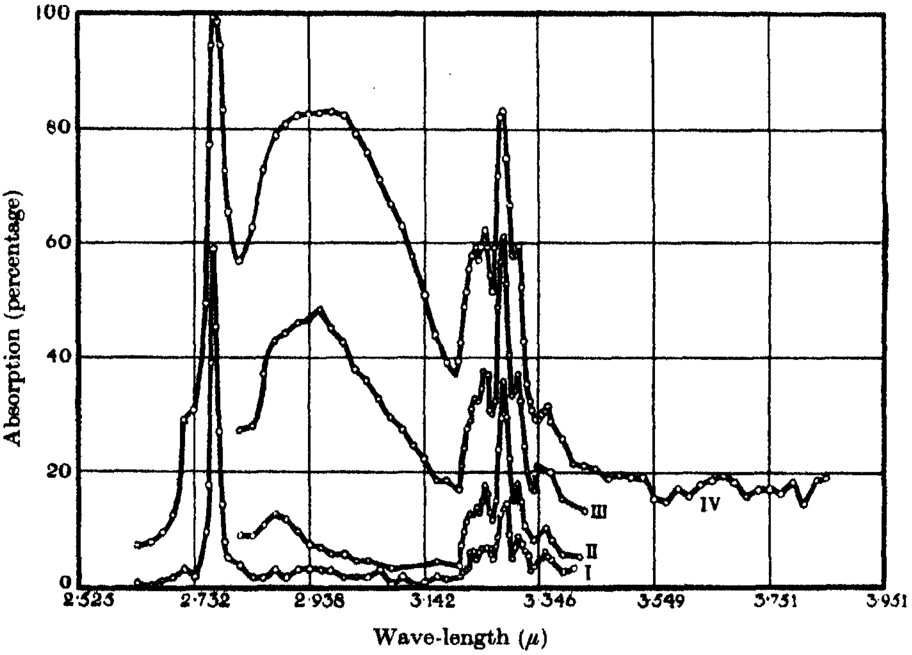


FIG. 1—Absorption of a 1·29 mm. layer of phenol in carbon tetrachloride: I, 0·0145; II, 0·0435; III, 0·087 and IV, 0·174 mol./l.

In fig. 1 the OH band at 2·769μ is affected to some extent at the higher concentration by overlap of the association band at 2·95μ, but by extrapolating this band over the region of the shorter wave-length band and deducting the appropriate value of *d* for the association band at a given wave-length from the total density, the density corresponding to the short wave OH absorption can be found. Fig. 2 provides an example of the method.

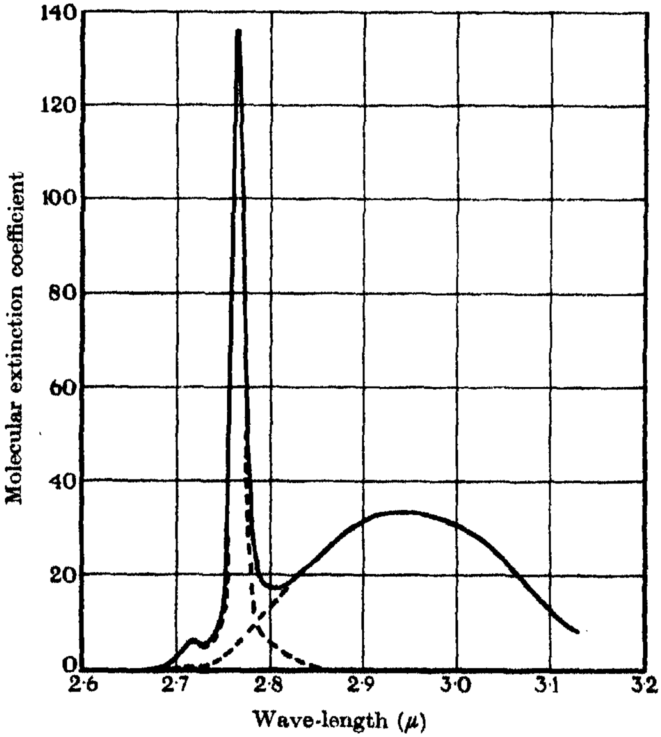


FIG. 2—The two OH bands of phenol showing their overlap and separation. *Solution of phenol in carbon tetrachloride, 0·174 mol./l.*

Table I shows the values of *d* over the OH band at 2·769μ for concentration (II) before and after elimination of the association band. The ratio of the density of the 2·769μ band to that obtained for the lowest concentration (I) is given for most of the band, and it will be seen that this ratio is nearly constant, the mean value being 2·60, although the variation of the concentration is actually 3 to 1. The difference between 3 and 2·60 represents those molecules which are linked up to give the association band. In a similar manner it is found that while the concentrations of (III) and (IV) are respectively 6 and 12 times that of (I), the intensity of the 2·769μ OH band is only 4·42 and 7·51 times that of (I). These results are set out in Table II. We prefer in the present instance not to measure the area under each absorption curve, since in regions away from the maximum some interference may be caused by other weak absorption bands, of which there is some evidence in the curves.

TABLE I—VARIATION WITH CONCENTRATION OF INTENSITY OF THE 2·769μ OH BAND FOR PHENOL IN CARBON TETRACHLORIDE. CELL LENGTH 1·29 MM.

Wave-length μ	Density of 2·769μ OH band (I) d_1 (0·0145 mol./l.)	Density of 2·769μ OH band (II) d_2 (0·0435 mol./l.)	Density of "association" band (II), d for μ of col. I	$\frac{d_2-d}{d_1}$
2·753	—	0·0879	0·0132	—
2·758	0·0793	0·226	0·0143	2·68
2·763	0·211	0·575	0·0154	2·66
2·768	0·380	0·953	0·0166	2·46
2·773	0·260	0·713	0·0177	2·68
2·778	0·133	0·352	0·0188	2·50
2·784	0·061	0·179	0·0200	2·61
2·789	—	0·110	0·0211	—
2·794	—	0·0802	0·0223	—

TABLE II—VARIATION WITH CONCENTRATION OF THE TWO OH BANDS FOR PHENOL IN CARBON TETRACHLORIDE

Solution as numbered in fig. 1	(I)	(II)	(III)	(IV)
Relative concentration, c_r	1	3	6	12
Relative intensity of 2·769μ OH band, d_r	1	2·60	4·42	7·51
Relative concentration of associated molecules not contributing to 2·769μ OH band, $c_a = c_r - d_r$	0	0·40	1·58	4·49
Intensity of association band as measured by density at peak (about 2·95μ), d	0	0·0615	0·279	0·801
c_a/d	—	6·5	5·7	5·6
$(d_r)^2/c_a$	—	16·9	12·4	12·6

Although the concentration of associated molecules varies widely, the ratio c_a/d changes but little and indicates that the intensity of the association band is proportional to the number of molecules not contributing to the 2·769μ band. The molecular extinction coefficient for the unassociated molecules, calculated from data in Table I, is 208 (cf. 2·7 found by Wulf and Liddel (1935) for the first harmonic) as against an average value of 1 for the associated molecules. The difference between these two OH values is, however, more than compensated by the much greater breadth and corresponding area of the association band. The approximate constancy of the expression $(d_r)^2/c_a$ in Table II would suggest that an equilibrium is set up between single and double phenol molecules in the concentration range studied: $PhOH + PhOH \rightleftharpoons 2PhOH$. Since d_r is proportional to the concentration of single molecules and c_a is proportional to the concentration of double molecules, $(dr)^2/c_a$ should be constant, according to the law of mass action. The high value 16·9 shown for concentration (II) probably arises from the difficulty in obtaining an accurate value of c_a when it is a small difference. The manner in which the phenol molecules are linked together is not well understood, and we shall consider this matter in detail later in this communication (p. 433).

TABLE III—WAVE-LENGTHS AND MOLECULAR EXTINCTION COEFFICIENTS FOR THE OH BANDS OF VARIOUS HYDROXY COMPOUNDS IN SOLUTION

			Sharp OH band		Association band	
Substance	Solvent	Conc. mol./l.	λ max (μ)	κ max	λ max (μ)	κ max
Phenol	CCl_4	0·174	2·771*	130	2·95	33·8
		0·087	2·769	154	2·94	23·8
		0·0435	2·769	181	2·89	11·0
		0·0145	2·769	208	—	—
	CCl_4 + 17% C_2Cl_6	0·1	2·771	141	2·92	19·5
	C_2Cl_4	0·1	2·771	139	2·88, 2·98	20·1, 17·2
	$CHCl_3$	0·1	2·782*	245*	2·93	14·2
Benzyl alcohol	CCl_4	0·169	2·764	34	2·97	28·5
	CCl_4	0·05	2·763*	58	2·87	8·2
	$CHCl_3$	0·1	2·770	58*	2·89	8·8
Diphenyl-carbinol	CCl_4	0·1	2·765	63*	2·89	13·5
	CCl_4	0·05	2·767*	60*	2·88	7·3
	$CHCl_3$	0·1	2·776	77*	2·89	7·5
Triphenyl-carbinol	CCl_4	0·05	2·769	75	2·87	3·5
	$CHCl_3$	0·1	2·781	61*	—	—
n-Butyl alcohol	CCl_4	0·1	2·750	41	2·89, 2·98	12·7, 14·9
tert-Butyl alcohol		0·1	2·765	41	2·87, 2·98	11·1, 10·0
Stearyl alcohol		0·1	2·750	45	2·88, 2·97	12·7, 14·0
Cetyl alcohol		0·112 (conc. by wt. equal to stearyl alcohol)	2·750	41	2·87, 2·99	14·6, 17·6

In the Table * denotes a somewhat lower accuracy for the value so marked.

Phenol and some Alcohols in various Solvents—In addition to carbon tetrachloride the following solvents were used for phenol: chloroform, tetrachloroethylene, and a saturated solution of hexachloroethane in carbon tetrachloride. Aliphatic alcohols were examined in carbon tetrachloride, and aromatic alcohols in carbon tetrachloride and chloroform solutions. The collected results including those for phenol in carbon tetrachloride are given in Table III (see figs. 3, 4 and 5). Wave-lengths and molecular extinction coefficients are given for both OH bands. It will be observed that the results for phenol are very similar for the non-polar solvents, but that the association is considerably reduced in chloroform. This latter solvent has the disadvantage that it has strong absorption in the region of 3μ but it dissolves some hydroxy compounds, e.g. triphenyl-carbinol, more readily than carbon tetrachloride.

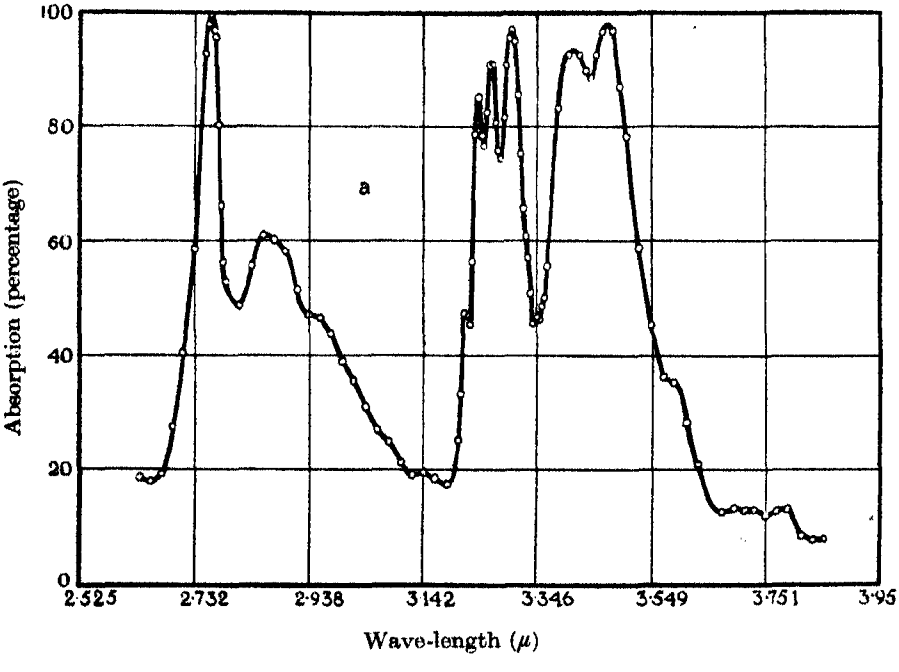


FIG. 3 (*a*)

FIG. 3—Absorption of aromatic alcohols in carbon tetrachloride (0·05 mol./l., 10 mm. layer). (*a*) Benzyl alcohol, (*b*) diphenyl-carbinol [(*b'*) 0·1 mol./l., 1·29 mm. layer], (*c*) triphenyl-carbinol [(*c'*) 0·066 mol./l., 1·29 mm. layer].

It is interesting to compare the results with some obtained by Errera and Mollet for ethyl alcohol. These are given displayed in Table IV.

The extrapolated value of κ max for completely associated EtOH is 120 compared with 91 calculated for phenol from data in Table II. A value of this order is also obtained for liquid water from the results of Ganz (1937).

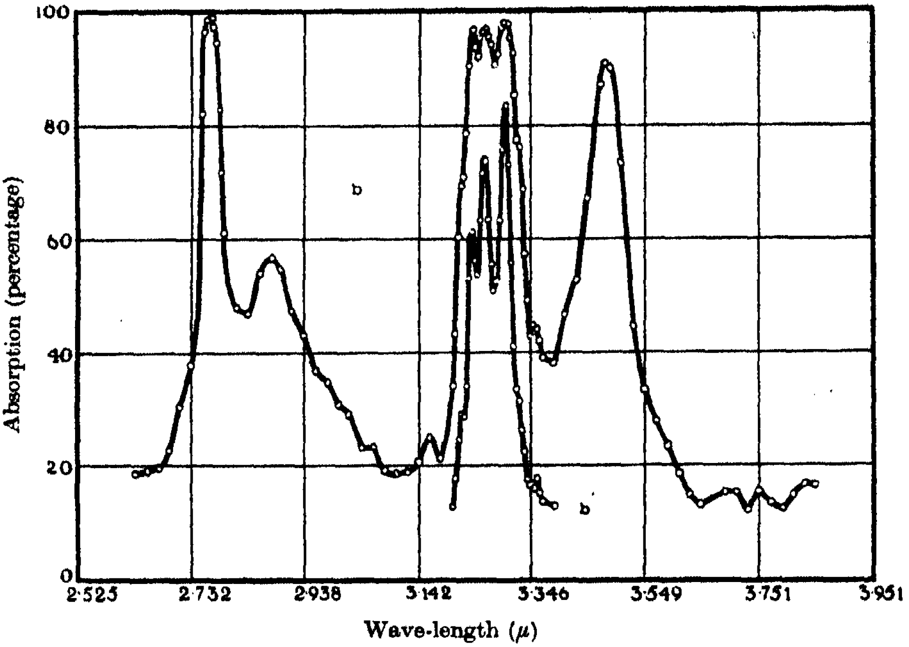


FIG. 3 (*b*), (*b*′)

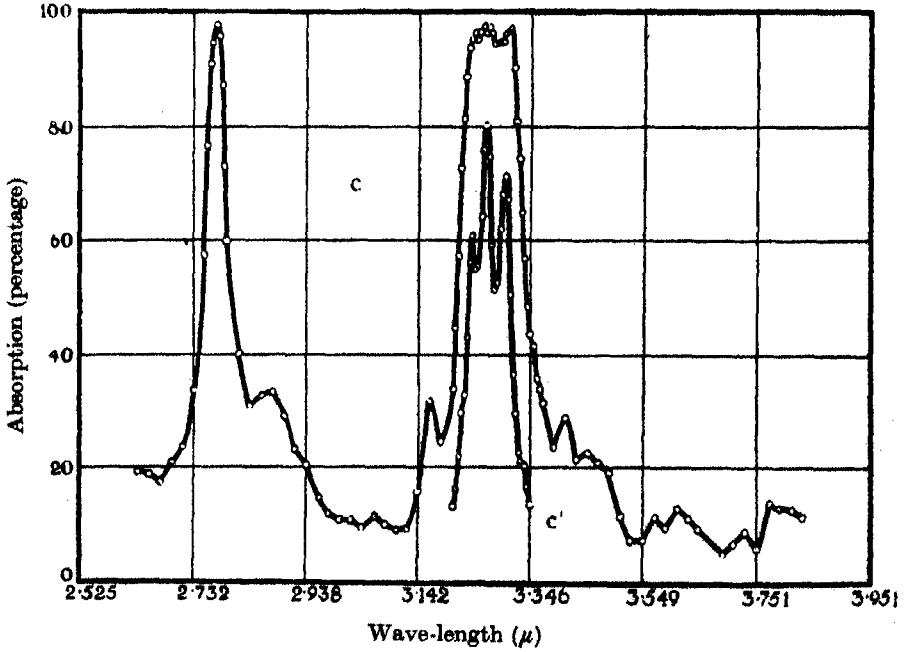


FIG. 3 (*c*), (*c*′)

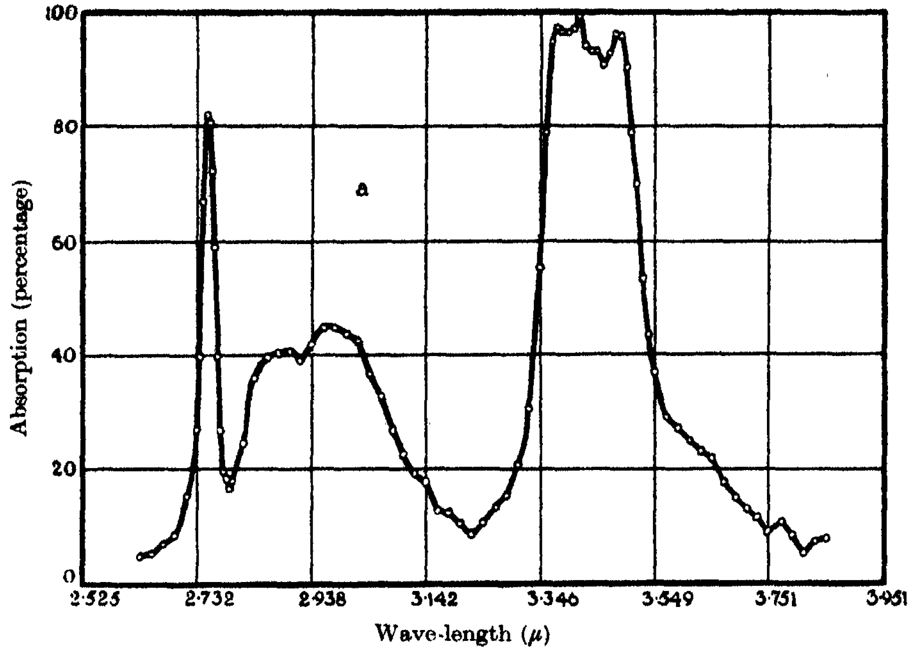


FIG. 4 (*a*)

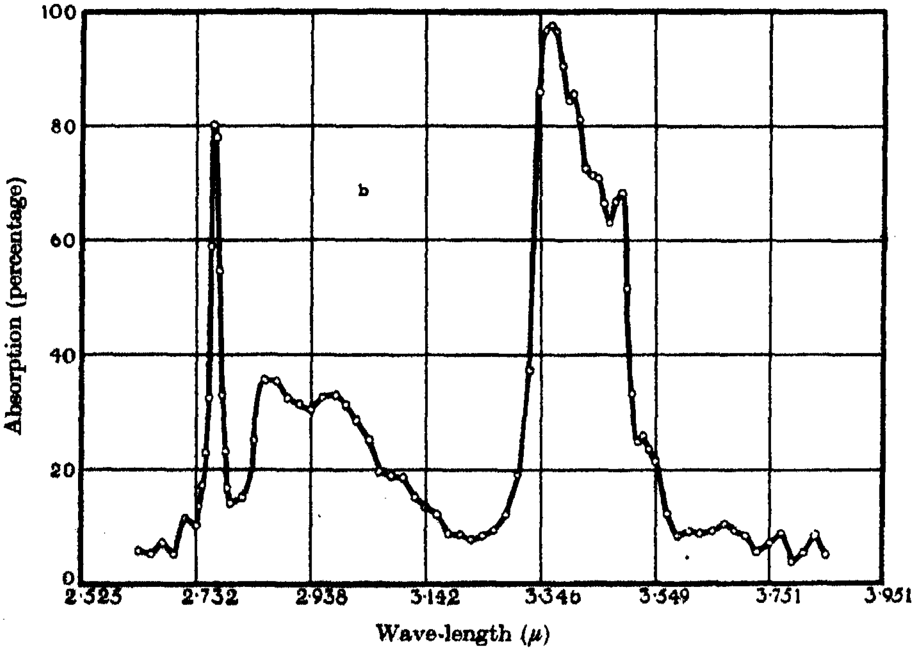


FIG. 4 (*b*)

FIG. 4—Absorption of aliphatic alcohols in carbon tetrachloride (1·74 mm. layer). (*a*) *n*-Butyl alcohol, (*b*) *tert*-butyl alcohol, (*d*) stearyl alcohol, all 0·1 mol./l.; (*c*) cetyl alcohol (0·112 mol./l.).

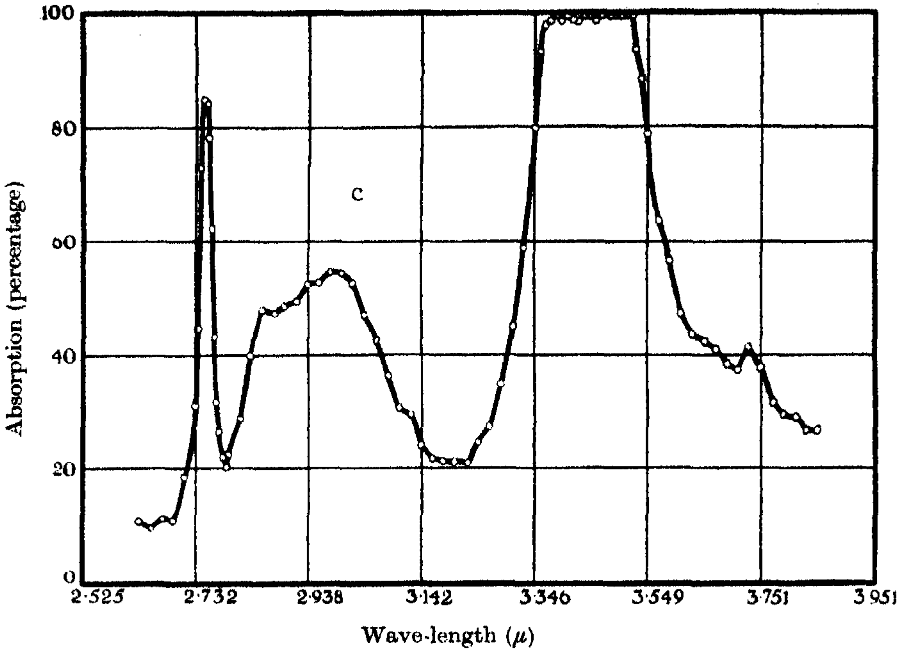

100
80
60
40
20
0
Absorption (percentage)
c
2·525
2·732
2·938
3·142
3·346
3·549
3·751
3·951
Wave-length (μ)

FIG. 4 (*c*)

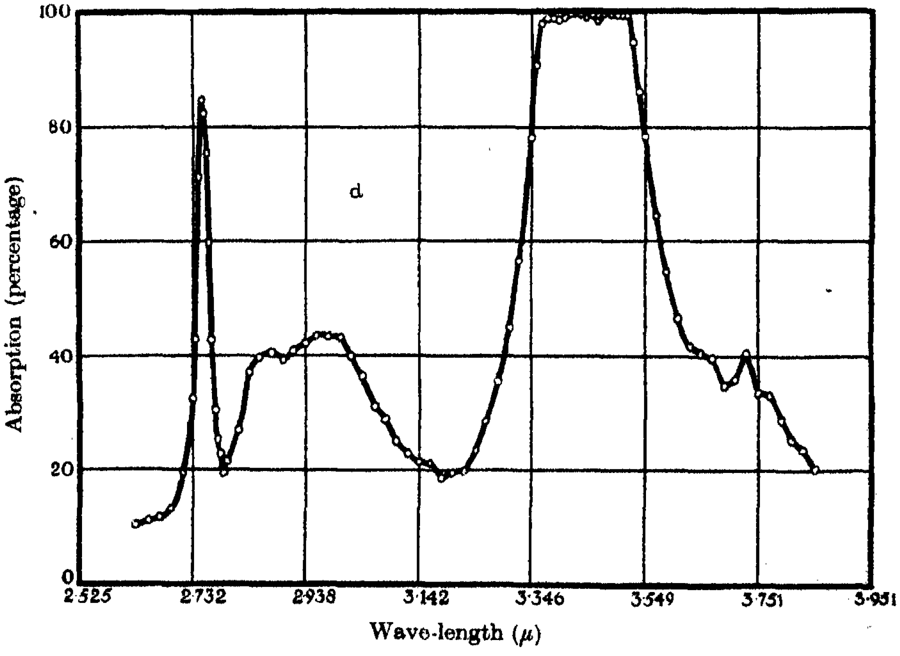

100
80
60
40
20
0
Absorption (percentage)
d
2·525
2·732
2·938
3·142
3·346
3·549
3·751
3·951
Wave-length (μ)

FIG. 4 (*d*)

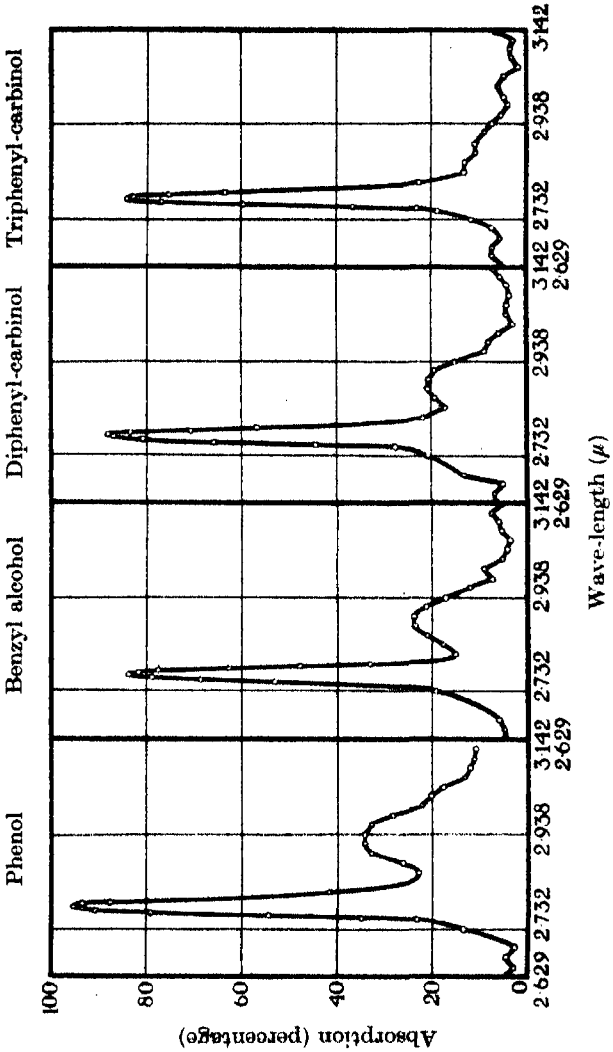


Fig. 5—Absorption of phenol and aromatic alcohols in chloroform (0·1 mol./l., 1·29 mm. layer).

In this case the association is complete, no sign of the shorter wave-length OH band being detectable.

Table V contains the frequency ν_1 of the short wave OH band found by us for several compounds, the first overtone ν_2 (Wulf and others 1935, 1936) and in one or two cases the second and third overtones ν_3 and ν_4 obtained by Mizushima, Uehara and Morino (1937). In this Table we have calculated the anharmonicity factor for the OH vibrations and hence the approximate energy D involved in the dissociation of the OH linkage.

TABLE IV—INTENSITY OF OH ASSOCIATION BAND OF ETHYL ALCOHOL CALCULATED FROM THE DATA OF ERRERA AND MOLLET (1937) FOR SOLUTIONS IN CARBON TETRACHLORIDE AND FOR THE VAPOUR

State	Concentration mol./l.	Association band λ max (μ)	Association band κ max
CCl_4 solution	3·33	2·97	105
	0·33	2·97	66
	0·10	2·965	25·2
Vapour 220 mm. Hg, 80° C.	0·01	2·88, 2·96	0·19, 0·19

TABLE V—FUNDAMENTAL FREQUENCIES AND OVERTONES FOR HYDROXY COMPOUNDS IN CARBON TETRACHLORIDE

Substance	Frequency (cm.$^{-1}$)	$x\omega$	ω	D (mean) K cal./mol.
Phenol	ν_1 3611	86	3783	115
	ν_2 7050	—	—	
	ν_3 10332	81	3768	
	ν_4 13430	87	3790	
n-Butyl alcohol	ν_1 3636	—	—	107
	ν_2 —	84	3804	
	ν_3 10406	—	—	
	ν_4 13459	104	3885	
Triphenyl-carbinol	ν_1 3611 ν_2 7050	86	3783	113
o-Phenyl-phenol*	ν_1 3607 ν_2 7035	90	3786	106
	ν_1 3564 ν_2 6945	92	3747	
Water vapour, included for comparison (Sleator and Phelps 1925)	ν_1 3742 ν_2 7235	125	3991	85

* This compound is not discussed in this paper, but is included here as it provides an interesting example of the relation between the frequency of a fundamental and that of its first overtone.

The fundamental and harmonics are here related by the simple expression $\nu_n = n\omega[1-(n+1)x]$, where ω is the natural frequency for very small displacements and x is the anharmonicity constant. The energy of dissociation D of the OH linkage, after allowing for zero-point energy, is $\dfrac{h(\omega - x\omega)^2}{4x\omega}$ or $\dfrac{h\omega(1-x)^2}{4x}$. As it will be necessary to link up the lowering of frequency of the

OH band, when association occurs, with the energy involved, we may here consider the magnitude of the changes resulting from the formation of an inter- or intramolecular hydrogen bond. Pauling (1936) has calculated the energy of the intramolecular hydrogen bond in *o*-chlorophenol, which causes the first harmonic of the OH band to shift from 7050 to 6910 cm.$^{-1}$, and obtains the value 1400 cal./mol. If we make the assumption that the anharmonicity factor, x, remains the same for both bands, then we obtain from the expression given above $\delta D/D = \delta\omega/\omega$, and this leads to a value about 2000 cal./mol. for *o*-chlorophenol, which is not very different from Pauling's value. Accordingly, for the association band which is found in phenol and alcohols at a frequency 150–300 cm.$^{-1}$ lower than the OH band for unassociated molecules, the energy involved will be between 3000 and 6000 cal./mol.

CH *bands*

Phenol and Alcohols in Carbon Tetrachloride. Figs. 1 and 3 show examples of the characteristic absorption due to aromatic CH groups in the region about 3·27μ. Generally speaking, this band consists of three sharp and intense absorption peaks which overlap, together with one or two weaker bands on

Table VI—Wave-Lengths and "Molecular" Extinction Coefficients (per CH) for the Aromatic and Aliphatic CH Bands of various Hydroxy Compounds in Carbon Tetrachloride

Aromatic CH

Benzene*		Phenol		Benzyl alcohol		Diphenyl-carbinol		Triphenyl-carbinol	
λ max (μ)	κ max	λ max (μ)	κ max	λ max (μ)	κ max	λ max (μ)	κ max	λ max (μ)	κ max
—	—	—	—	3·214	0·62	3·217	0·49	—	—
3·232	7·1	3·234	1·8	3·237	2·7	3·237	2·3	3·238	1·9
		3·249	1·8						
3·253	4·8			3·261	4·6	3·263	3·8	3·265	4·7
		3·283	6·2						
3·291	9·5			3·298	6·4	3·300	5·5	3·300	3·5
		3·311	1·3						

* The values for benzene determined by us are included for comparison and apply to a solution in carbon tetrachloride.

Aliphatic CH

Diphenyl-carbinol		Benzyl alcohol		*Tertiary*-Butyl alcohol	
λ max (μ)	κ max	λ max (μ)	κ max	λ max (μ)	κ max
—	—	—	—	3·36	9·7
—	—	3·40	8·9	3·40	ca. 5
3·479	19·4	3·477	12·6	3·48	2·7

the sides of the main band. The wave-lengths and approximate "molecular" extinction coefficients (calculated for one aromatic CH group) are given in Table VI.

The aliphatic CH bands (figs. 3 and 4) are similarly presented. For the purpose of estimating the "molecular" extinction coefficient of the CH group, CH_2 is considered as providing two and CH_3 three CH linkages. Triphenyl-carbinol with no aliphatic CH linkage has of course no important absorption in the region of $3{\cdot}48\mu$.

DISCUSSION OF RESULTS

OH *bands*—The results obtained in the 3μ region of the spectrum show a close correspondence with the data given by Wulf and others for the $1{\cdot}5\mu$ region as regards the OH absorption band. In many cases our curves agree closely with those of Wulf in spite of the fact that the intensity of the first overtone is only a few per cent of the fundamental in the 3μ region.

We find, as does Wulf, that the position of the OH band for the unassociated molecules varies a little from one compound to another. We summarize our results for solutions in carbon tetrachloride as follows:

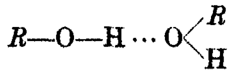
Substance	λ max (μ)
n-Aliphatic alcohols	2·750
tert-Butyl alcohol	2·765
Benzyl alcohol	2·764
Diphenyl-carbinol	2·765
Triphenyl-carbinol	2·769
Phenol	2·769

It is of interest to recall the value for water vapour, $2{\cdot}672\mu$. *n*-Aliphatic alcohols, even with long chains, have a lower wave-length than benzyl alcohol which also contains the $—CH_2OH$ group. *tert*-Butyl alcohol similarly has a somewhat lower wave-length than the tertiary aromatic alcohol, triphenyl-carbinol. Phenol with the group $>$COH has the same wave-length as triphenyl-carbinol, while diphenyl-carbinol with the group —CHOH has a slightly lower wave-length. We thus see that when in an aliphatic or aromatic hydroxy compound a C, CH or CH_2 group is directly attached to the OH, the frequency of this group is not independent of the nature of the rest of the molecule. This may well have some connexion with the fact that the aromatic-aliphatic C—C distance in compounds such as methyl benzenes, as found by Robertson (1933) from X-ray, and by Jones (1935) from electron diffraction measurements, is only 1·47 or 1·50 A instead of the 1·54 A for the C—C distance in saturated aliphatic compounds.

Our results for phenol show how effective a given OH unit is in absorbing radiation, according to whether it exists in an isolated molecule or in an association complex. It becomes possible to determine the proportions of single and complex molecules in a solution from the absorption data. A further point that emerges from the results of Errera and Mollet on ethyl alcohol, Ganz on liquid water, and ours on phenol is that when associated, a given OH unit is equally effective in absorbing radiation over a wide range of concentration, from a dilute solution in carbon tetrachloride to the pure liquid, although the frequency of the position of maximum absorption increases as the concentration is reduced. Ganz (1937) found little change in the intensity of the 3μ band of liquid water with increasing temperature (an observation which we have confirmed) and this is not surprising since the "concentration" is so great. The frequency of the position of maximum absorption of this band however does increase with rising temperature in accordance with some reduction in the strength of association.

From fig. 2 it will be noticed that the association band extends right up to the normal $2{\cdot}769\mu$ OH band, but not to higher frequencies; its maximum moves towards lower frequencies as its intensity becomes greater (fig. 1). It is considered that at any moment complex molecules exist in the solution with a strength of binding varying from the maximum value to nearly zero. As this binding becomes stronger so the OH linkage is weakened and the OH frequency is reduced (cf. Bernal 1937). The association band may be regarded as a kind of distribution curve showing the proportion of molecules with any given strength of binding.

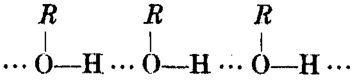
The association of molecules of general type ROH has been ascribed to the formation of hydrogen bonds, thus:



$$R\text{—O—H}\cdots\text{O}\begin{smallmatrix} R \\ \text{H} \end{smallmatrix} \tag{1}$$

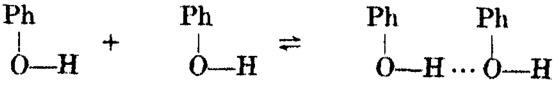
According to Bernal (1937) a distinction can be made between hydrogen bonds and hydroxyl bonds. In the hydrogen bond the hydrogen atom is attached with almost equal firmness to two oxygen atoms, the O—O distance being 2·5–2·65 A corresponding to an energy of the bond about 8000 cal./mol. The OH bonds are in these conditions greatly weakened and the characteristic OH band is not found in the normal spectral region. In the hydroxyl bond the hydrogen atom is still regarded as belonging to its parent oxygen, although the OH bond is somewhat weakened and the band is observed at a lower frequency (by some 150–300 cm.$^{-1}$). The distance O—H ⋯ O in this type of bond is taken to be 2·7–2·9 A and the energy involved about 5000 cal./mol. This view does not seem to be quite conclusive since the dissociation energy of a normal OH bond (106,000 cal./mol.) is very large compared with the changes in energy suggested for the hydrogen bond above. The "disappearance" of the normal OH frequency about 2·75μ when a definite hydrogen bond can be postulated may be due to the difficulty of locating or recognizing the shifted band. The frequency shift to be expected for 8000 cal./mol. corresponding to a strong hydrogen bond is about 400 cm.$^{-1}$ and this would displace the OH band from its position at 2·75μ to about 3·1μ. In this region of the spectrum other bands are found, arising for example from aromatic CH frequencies, and the resulting absorption curves are rather complex. It is significant, however, that we have found that strong absorption occurs in salicylic aldehyde and in fatty acids where strong hydrogen bonds exist (cf. Buswell, Deitz and Rodebush 1937).

On the supposition that as the concentration of hydroxy-body increases the first interaction is to give double molecules such as (1), it follows that when this type of association becomes complete the molecular extinction coefficient for the sharp OH band will fall to half the value found at infinite dilution, since only half the normal OH bonds giving rise to this frequency now remain. This effect, if it exists, will be masked at higher concentrations by the formation of larger complexes with a still smaller number of "free" OH linkages:



$$\cdots\overset{R}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H}\cdots\overset{R}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H}\cdots\overset{R}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H}\cdots$$

The formation of double molecules of type (1) does not accord with the observed variation of the intensity of the two bands with concentration in the case of phenol. If we suppose the complexes to be in equilibrium with single molecules



$$\overset{\mathrm{Ph}}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H} + \overset{\mathrm{Ph}}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H} \rightleftharpoons \overset{\mathrm{Ph}}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H}\cdots\overset{\mathrm{Ph}}{\overset{|}{\mathrm{O}}}\!-\!\mathrm{H} \qquad (2)$$

we shall have, adopting our previous notation: total number of phenol molecules c_r, number of phenol molecules in association c_a, number of single phenol molecules $c_r - c_a$, number of "free" OH linkages $c_r - \frac{c_a}{2}$. The difference $c_r - d_r$ will now become $\frac{c_a}{2}$ instead of c_a. We give values of c_a on this basis together with $(c_r - c_a)$ and $\frac{(c_r - c_a)^2}{c_a}$.

c_r	1	3	6	12
c_a	0	0·80	3·16	8·98
$c_r - c_a$	1	2·20	2·84	3·02
$\frac{(c_r - c_a)^2}{c_a}$	—	6·05	2·55	1·02

The value $\frac{(c_r - c_a)^2}{c_a}$ should be constant if an equilibrium such as (2) is set up. Actually the variation in this quantity is seen to be far greater than the corresponding variation in concentration. This difficulty would be removed if the "free" OH in the double molecule also changes its frequency to the region of the association band, but this introduces an artificial distinction between the two hydrogen atoms. In our view the observations here recorded suggest interaction between any two molecules in such a way that both OH bonds are affected, the magnitude of the effect depending on the relative position and distance apart of the molecules.

It is of some interest to consider the mutual interaction of two molecules. London (1937) gives three constituents of van der Waals force. These, expressed in terms of energy of interaction, may for the present purpose be given briefly as follows:

1—Orientation effect due to the mutual interaction of permanent dipoles. Both attractive and repulsive forces arise, but averaging for all possible orientations the mean energy for the interaction between two dipoles with the same moments μ is given by $\overline{U} = -\frac{2}{3}\frac{\mu^4}{r^6 kT}$, the negative sign implying attraction. r is the distance between the dipoles, k is Boltzmann's constant, and T the absolute temperature. For low temperatures or small distances $\left(kT \lesssim \frac{\mu^2}{r^3}\right)$ the expression becomes invalid, since it leads to energies of interaction greater than the maximum possible, $-\frac{2\mu^2}{r^3}$, i.e. the interaction energy for two molecules oriented in the most favourable manner possible, with their dipoles parallel to each other along the line joining them.

2—Induction effect due to the interaction of induced dipole and permanent dipole responsible for it. The force between the molecules is always attractive and the energy is very much less than that in 1 for the type of compound here considered.

3—Dispersion effects due in a simple case to the interaction of two three-dimensional isotropic harmonic oscillators with no permanent moment in their rest position. During oscillations various configurations of nuclei and electrons occur, and in general these are equivalent to rapidly varying dipoles which induce dipoles in other molecules and so produce a varying energy of interaction which in the simple case considered is $\overline{U} = -\frac{3}{4}\frac{h\nu_0\alpha^2}{r^6}$ (attraction), where α is the polarizability.

London's calculations for H_2O show that the relative magnitudes of these forces in the order 1, 2 and 3 are in the ratio 190:10:47. In the case of phenol and alcohols the interactions we are concerned with are between OH groups, or better —COH groups. These we may in the first instance represent as spheres (see fig. 6) containing a —CO dipole and an —OH dipole which are added vectorially to give the resultant dipole, for simplicity supposed to behave as though at the centre of a sphere. Where α, the polarizability, is involved, we may consider a value appropriate for a —COH "molecule", since the energy of interaction falls off according to a high power of the distance, and more remote portions of the molecule, e.g. in triphenyl-carbinol, will only contribute slightly to the total energy. By doing this we can assume that the induction and dispersive effects will be much the same for all alcohols, being however somewhat reduced with big molecules owing to the difficulty of approach of the —COH groups of different molecules. We find, if r is taken as 2·7 A, the distance at which the hydroxyl bond forms, that the maximum possible orientation energy (for $\mu = 1{\cdot}7 \times 10^{-18}$ e.s.u.) is nearly 4·3 kcal./mol. To this value there should be added energy arising from the dispersion effect (London 1937, p. 19) which, for the hydroxy bodies here discussed we shall consider to be of the same order as for water. In that case the value should be increased by some 25 % and the total energy is of the right order of magnitude. Approach to this maximum value will become less likely as the temperature rises and increased thermal agitation tends to promote random distribution of the dipoles. Accordingly association should decrease with rise of temperature, as is well known to be the case. Dilution also reduces association, and this may be partly attributed to the effect of a large preponderance of solvent molecules hindering the favourable orientations of the dipoles which lead to

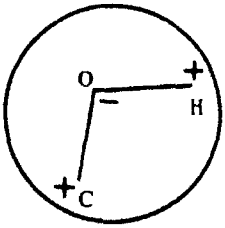


FIG. 6

appreciable interaction energy. Owing to the disturbing influence of the solvent molecules, two dipoles will only associate if they happen to come within a certain minimum distance from one another, r. If the concentration (expressed as n molecule of absorbing material per c.c.) is small, the probability that any molecule of absorbing substance will be within the distance r of a given molecule is $\frac{4}{3}\pi r^3 n$, i.e. $\propto n$, and since the total number of molecules subjected to the radiation is also proportional to n, the intensity of absorption for associated molecules will vary as n^2, a result which agrees with the equilibrium $R\mathrm{OH} + R\mathrm{OH} \rightleftharpoons 2R\mathrm{OH}$ for small concentrations (see p. 424).

The results in Table III show that changing the solvent from carbon tetrachloride to chloroform reduces the strength of the association band at 0·1 mol./l. in the ratios: phenol 1·83; diphenyl-carbinol 1·80; benzyl alcohol 2·38. We can thus state in a general way that the association is reduced by a factor of 2, and this is nearly the ratio of the dielectric constants of the two solvents (CCl_4 2·24; $CHCl_3$ 5·0). It will have been noticed that in the calculations made by London, the dielectric constant is taken as unity and this is reasonable if we consider single molecules isolated *in vacuo*. If, however, we regard the solvent as a continuous medium with its appropriate dielectric constant ϵ, then the force of attraction or repulsion between two dipoles and also the energy of interaction between them is proportional to $1/\epsilon$, and this may account for the observed reduction of association in changing from carbon tetrachloride to chloroform. An alternative explanation may be found in the fact that while carbon tetrachloride is non-polar, chloroform has a dipole moment of $1 \cdot 1 \times 10^{-18}$ e.s.u. The chloroform molecules may now be regarded as forming loose complexes with the solute molecules and thereby hinder hydroxyl association by reason of the difficulty of approach to the required distance 2·7 A. A small increase in this distance reduces the energy of interaction very considerably (up to the sixth power of the distance). It will also be observed from Table III that the normal OH wave-length varies slightly according to the solvent, it being higher in chloroform. These wave-length changes are: benzyl alcohol 0·006μ; diphenyl-carbinol 0·011μ; triphenyl-carbinol 0·012μ; phenol 0·013μ. This increase of wave-length may arise from the formation of a weak complex between the alcohol and chloroform molecules, for association causes a reduction of frequency of the OH bond. The average displacement of 0·01μ = 13 cm.$^{-1}$ is equivalent to about 260 cal./mol. (p. 431).

It is of interest to note that increasing molecular complexity generally reduces the association both in carbon tetrachloride and chloroform, as is seen from the series PhOH, $PhCH_2OH$, Ph_2CHOH and Ph_3COH (figs. 3

and 5), the association being greatest for phenol and least for triphenylcarbinol. *n*-Butyl alcohol and stearyl alcohol, however, are equally associated, while *tertiary*-butyl alcohol is less and cetyl alcohol a little more associated than the two former. The frequency shift between the sharp OH band and the centre of the association band for concentrations about 0·1 mol./l. is 220 cm.$^{-1}$ and this is equal to 4000 cal./mol. or more, an energy considerably greater than that involved in many cases for intramolecular bonding, e.g. 1400 cal./mol. for *o*-chlorophenol.

Measurements of the heats of dilution of alcohols give similar values for the energy of interaction and indicate the possibility of more complex groupings than double molecules (Wolf 1937). It will be observed that the association band, particularly for aliphatic alcohols (fig. 4), has at least two maxima, and it may be of significance in this connexion that the separation of these maxima (the mean positions for which are 2·88 and 2·98μ for all the hydroxy bodies studied) is almost that between the sharp OH band and the first maximum at 2·88μ. The apparent shift towards longer wave-lengths of the association band with increase of concentration may find its explanation by supposing that the band at 2·88μ which is more pronounced at lower concentrations is due to double molecules, while the second longer wave-length band arises from more complex combinations and therefore only becomes marked at higher concentrations. The reduction of association in either solvent with increasing molecular complexity can be attributed to the reduction of the number of orientations of the polar —COH groups for which interaction can occur. Particularly pronounced is this effect in triphenylcarbinol in which nearly half the space surrounding the —COH group, within the critical distance 2·7 A, is not accessible to another —COH group (fig. 7).

CH *bands*—From figs. 1, 3 and 4, and Table VI, the differing behaviour in absorption of the aliphatic and aromatic CH groups is brought out. The single aliphatic CH group in diphenyl-carbinol gives rise to one strong absorption band at 3·479μ, whereas the CH_2 group in benzyl alcohol gives one band at 3·477μ and a slightly weaker one at 3·40μ. Fig. 4 (*c*) and (*d*) for cetyl- and stearyl-alcohols show very intense CH absorption of the CH_2 group in the same position as in benzyl alcohol. The curve for *tert*-butyl alcohol, fig. 4 (*b*), shows the corresponding absorption for the CH_3 group and this is even more complex than for CH_2. Three distinct components can be identified at 3·48, 3·40 and at 3·36μ, the last being the most intense. The total value of κ max per aliphatic CH in Table VI is about 20, and this is divided between the two bands for CH_2 and the three (or more) bands for CH_3. The value of κ max for the 3·40μ band of *tert*-butyl alcohol is difficult to assess from the curve; it is, however, less than the value 9·7 for the 3·36μ band.

The aromatic CH bands do not vary much in position in the spectrum in the series benzene, benzyl alcohol, diphenyl-carbinol, and triphenyl-carbinol, although there is a slight progressive increase of wave-length with increasing molecular complexity. It should be noted, however, that the three strong bands do not always retain the same relative intensities (cf. Table VI and figs. 1 and 3). The aromatic CH bands in phenol are very different from those of the aromatic alcohols, one band being particularly strong at the expense of the others. This peculiarity of phenol appears to be due to the direct attachment of the OH group to the benzene ring and disappears when another group, such as CH_2 is interposed. The variations which occur in the aromatic CH bands of different hydroxy-compounds are valuable as a means of identification when considered in conjunction with the characteristics of the OH group itself, e.g. phenol and triphenyl-carbinol possess OH bands at exactly the same wave-length, but the CH bands are somewhat different as regards intensity and position, although the bands of both compounds are easily recognized as being due to aromatic CH groups.

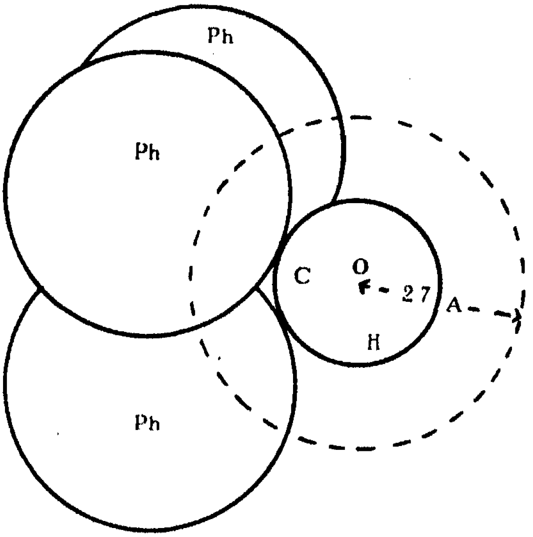


FIG. 7

We desire to thank Mr. F. S. Benge of this laboratory for his valuable assistance in constructing apparatus and for assisting in taking observations.

SUMMARY

1—The two OH absorption bands in the region of 3μ for phenol and certain alcohols have been investigated for solutions in carbon tetrachloride and other solvents. With phenol the intensity of the longer wave "association band" is proportional to the number of molecules not contributing to the short wave band. The association band is complex and its maximum moves to shorter wave-lengths as its intensity falls. At moderate concentrations an equilibrium is set up between single and double phenol molecules, and the forces acting between such molecules are considered in the light of London's theory of intermolecular forces. In chloroform as a solvent the intensity of the association band is reduced approximately one-half compared with the solution in carbon tetrachloride; increasing molecular complexity in the order PhOH, $PhCH_2OH$, Ph_2CHOH and Ph_3COH reduces association in either solvent.

Both long chain normal alcohols examined in carbon tetrachloride solution show undiminished association. Increasing molecular complexity in the series $PhCH_2OH$, Ph_2CHOH and Ph_3COH causes a small progressive increase in the wave-length of the short wave OH band; and when carbon tetrachloride is replaced by chloroform as a solvent, this band is displaced to a slightly longer wave-length. The structure of the molecule itself may considerably affect the magnitude of the association band, e.g. the presence of three phenyl groups in triphenyl-carbinol almost suppresses association.

2—Aliphatic CH bands for CH, CH_2 and CH_3 groups are examined and it is shown that in the cases considered there is a constant intensity of absorption for the CH linkage in each molecule, but that it is distributed among the one, two or three components attributed to each group respectively.

Aromatic CH bands retain their positions, but not necessarily their relative intensities, for a series of aromatic alcohols; with phenol, bands of somewhat different position and intensity are found, this being attributed to the direct attachment of the OH group to the benzene ring.

REFERENCES

Barchewitz, P. 1937 *C.R. Acad. Sci., Paris*, **204**, 1184.
Bernal, J. D. 1937 *Trans. Faraday Soc.* **33**, 142 and 210.
Buswell, A. M., Deitz, V. and Rodebush, W. H. 1937 *J. Chem. Phys.* **5**, 84.
Errera, J. 1937 *Trans. Faraday Soc.* **33**, 120.
Errera, J. and Mollet, P. 1936 *Nature, Lond.*, **138**, 882.
—— —— 1937 *C.R. Acad. Sci., Paris*, **204**, 259.
Freymann, R. 1937 *C.R. Acad. Sci., Paris*, **204**, 1063.

Ganz, E. 1937 *Ann. Phys., Lpz.* [v], **28**, 445.
Jones, P. L. F. 1935 *Trans. Faraday Soc.* **31**, 1036.
London, F. 1937 *Trans. Faraday Soc.* **33**, 8.
Mizushima, S., Uehara, Y. and Morino, Y. 1937 *Bull. Chem. Soc. Japan*, **12**, 132.
Pauling, L. 1936 *J. Amer. Chem. Soc.* **58**, 94.
Robertson, J. M. 1933 *Proc. Roy. Soc.* A, **142**, 659.
Robertson, R., Fox, J. J. and Martin, A. E. 1934 *Philos. Trans.* A, **232**, 463.
Sleator, W. W. and Phelps, E. R. 1925 *Astrophys. J.* **62**, 28.
Wolf, K. L. 1937 *Trans. Faraday Soc.* **33**, 179.
Wulf, O. R. and Liddel, U. 1935 *J. Amer. Chem. Soc.* **57**, 1464.
Wulf, O. R., Liddel, U. and Hendricks, S. B. 1936 *J. Amer. Chem. Soc.* **58**, 2287.

The Influence of Temperature on the Capture of Neutrons by Various Elements

By J. L. Michiels and J. A. Saxton

(*Communicated by G. P. Thomson, F.R.S.—Received* 2 *July* 1937)

The velocity distribution of neutrons of thermal energies can be changed by passing the neutrons through a layer of paraffin wax about 1 cm. thick, maintained at a different temperature from that of the neutrons entering the layer (Moon and Tillman 1936). When neutrons of an energy corresponding to room temperature (290° K.) are passed through such a wax layer at 90° K., the radioactivity induced in elements by the absorption of these cooled neutrons is greater than if the layer is also at 290° K. For silver a wax layer of about 1 cm. thickness gives the greatest increase, about 35 %. The effect was investigated for several elements and varied from about 20 % for iodine to about 50 % for copper (about 30 % being the average order of magnitude of the effect).

The differences between these values may arise from two causes. First, an element may absorb, selectively, neutrons of an energy far above thermal energy. The activity due to these neutrons will be independent of temperature, and the measured value of the temperature effect will be less than if the thermal neutrons alone had been concerned. Secondly, the variation of absorption probability with velocity may differ from one element to another in the "thermal region".

Since either an abnormally high or an abnormally low value of the temperature effect would indicate that the element in question deserved further investigation, it seemed desirable to extend measurements of these effects to elements not previously studied.

Moon and Tillman (1936), using temperatures of 290 and 90° K., measured temperature effects, corresponding to the geometrical disposition of their apparatus, for silver, copper, vanadium, iodine, dysprosium and rhodium. Our experiments have been made over the same range of temperature with sodium, gold, indium, arsenic, bromine (18 min. period), manganese, aluminium; silver was also used as a check, having previously been thoroughly investigated.

While the geometrical arrangements which we used were very similar to those used by Moon and Tillman much time was saved by irradiating two specimens simultaneously, one within a cold and the other within a warm wax cylinder. The activities of the two specimens were compared, and then (after this activity had thoroughly decayed) the experiment was repeated with the temperatures of the cylinders interchanged. If the specimens and cylinders are identical, and if they are symmetrically disposed with respect to the source, then the two ratios $\left(\frac{\text{Activity due to cold neutrons}}{\text{Activity due to warm neutrons}}\right)$ will be equal. All small errors due to asymmetry are accurately eliminated by taking the geometric mean of the two ratios.

A large rectangular block of paraffin wax had two symmetrically placed identical cylindrical cavities cut in it. Between these two, and equidistant from them, a smaller cavity was made to hold the source of neutrons—a radon-beryllium bulb, contained for convenience in a long glass tube (fig. 1). The source was usually equivalent to about 100 mC. of radium as measured by its γ-ray activity.

In the two cylindrical cavities were placed two similar vacuum flasks containing cylindrical layers of wax, the two layers being as nearly equal as possible. Two similar specimens were made for each of the elements investigated. The activity produced in a specimen irradiated by thermal neutrons through a layer of wax, the temperature of which can be changed, depends upon the geometrical disposition of source, specimen and surrounding wax, the temperature and thickness of the wax layer, and upon the nature and thickness of the specimen. Considering only specimens of a given element which are thin compared with the half-value thickness of that element for the absorption of slow neutrons, let us now define "J" as the ratio of the activity produced in such a specimen by neutrons which have passed through a layer of wax of given thickness maintained at a temperature of 90° K., to that produced under similar conditions when the wax layer is at 290° K. This, of course, refers to the particular geometry of our apparatus described above. Suppose, that for each element, the two similar specimens be designated as *A* and *B*, and that the two flasks be called *X* and *Y*. For a given wax thickness in the flasks, to determine *J*, two experiments were performed. The geometrical disposition (e.g. *A* in *X* and *B* in *Y*) was identical in the two cases, but in Exp. 1, *X* was at 90° K. and *Y* at 290° K., whilst in the second experiment the temperatures were reversed. From the first experiment we have $J_1 =$ (activity of A/activity of B) and from the second $J_2 =$ (activity of B/activity of A), and it is easy to show that the true temperature effect *J* is given by $\sqrt{(J_1 J_2)}$.

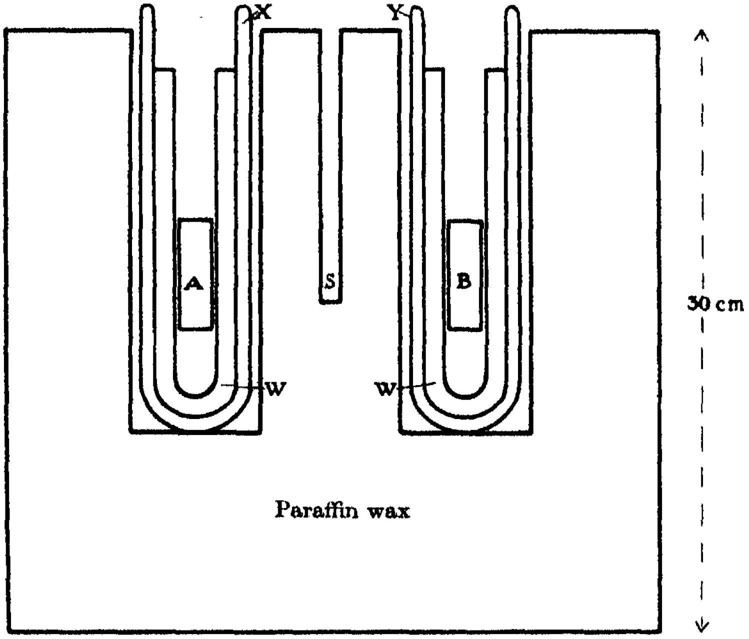


FIG. 1—*A*, *B*, specimens; *X*, *Y*, vacuum flasks; *W*, wax layer; *S*, neutron source.

The recorded activity of a specimen irradiated as described above will be some function of the form:

$$\text{Activity} = DFCST,$$

where the factor:

D depends on the specimen.

F depends on the flask containing specimen and its position.

C depends on the counter sensitivity.

S depends on the strength of the neutron source.

T depends on the temperature of the neutrons.

Now S is the same for A and B in any given irradiation; and, during the counting of the activity of the specimens in any one experiment C was assumed to be constant, this point is dealt with again below. Then we have in the first experiment:

$$\begin{aligned}\text{Activity of } A &= \alpha_1 = AXC_1S_1T_1,\\ \text{Activity of } B &= \beta_1 = BYC_1S_1T_2.\end{aligned}$$

And in the second experiment:

$$\begin{aligned}\text{Activity of } A &= \alpha_2 = AXC_2S_2T_2,\\ \text{Activity of } B &= \beta_2 = BYC_2S_2T_1.\end{aligned}$$

So that
$$\frac{\alpha_1}{\beta_1} = \frac{\alpha_2(T_1)^2}{\beta_2(T_2)},$$

therefore
$$J = \frac{T_1}{T_2} = \sqrt{\frac{\alpha_1\beta_2}{\beta_1\alpha_2}},$$

i.e.
$$J = \sqrt{J_1J_2},$$

since
$$\alpha_1/\beta_1 = J_1, \quad \beta_2/\alpha_2 = J_2.$$

The method described above is independent of small inequalities in the conditions of irradiation of the two specimens, and also of a change in counter sensitivity from the first experiment to the second. In order to correct for any possible small change of sensitivity in the counter—in the same direction, either increasing or decreasing—during the counting of the activity of the specimens in any one experiment, the specimens were placed round the counter in the order A, B, B, A; B, A, A, B, for equal intervals of time, usually two or three minutes. Several minutes observation of the natural count was made after each group of four counts of the specimens. The sensitivity of the counter was also checked from time to time by observing the rate of counting when a weak radio-thorium source was placed in a standard position with reference to the counter.

In general, sufficient counting was performed for a given value of J to have a probable error (not standard deviation) of about 2 % owing to the probable statistical fluctuations of the counts.

EXPERIMENTAL DETAILS

The flasks were of the "Isovac Steel Glass" variety, these being found most suitable for the purpose. They showed no measurable absorption of neutrons. Also several of the irradiations were of necessity long, since the elements in question had radioactive isotopes of long period; and it was found that a layer of wax inside one of these flasks, cooled down to the temperature of liquid air, retained its temperature for some considerable time. It was found necessary to recool the wax layer only every two hours or so during a long irradiation.

The wall separation of each flask used was examined over the whole of the flask, so that the wax layers in a given pair of flasks could be made of equal thickness. In all four pairs of flasks were used, with layers of 0·7, 0·9, 1·1 and 1·3 cm. thickness. After all observations had been made the flasks were broken and the wax examined for faults and inequalities in thickness.

Two similar specimens were made for each of the elements investigated. For sodium we first tried specimens made up of sodium fluoride, these were unsatisfactory, their activity being very weak. Much greater success was obtained using metallic sodium of thickness 0·055 g./cm.^{2} wrapped up in silver foil. For gold, indium, aluminium and silver, the metals were also used of thicknesses 0·16, 0·08, 0·15 and 0·08 g./cm.^{2} respectively. In the cases of bromine, manganese, arsenic, however, we had to use one of the more suitable of their respective compounds. Sodium bromide (0·15 g./cm.^{2}), manganese dioxide (0·2 g./cm.^{2}), arsenious oxide (0·15 g./cm.^{2}), were used. All the specimens were of a thickness small compared with the half-value thickness for the absorption of slow neutrons. These specimens involving the use of powders were made up in the following manner. The necessary amount of the compound in question was included between two pieces of thin celluloid, bent into cylindrical form, rigidity being obtained by fixing narrow strips of nickel along the edges of the specimens. The powder was kept in position between the celluloid by a small pad of cotton-wool.

On absorbing neutrons the elements examined become β-radioactive, so that the activities of our specimens were determined by placing them round a Geiger-Müller counter connected to an amplifier and mechanical recording device.

RESULTS AND DISCUSSION

Values of J were obtained for each element for the four wax thicknesses in the flasks. The values obtained are given in Table I; the ratios have, in general, a probable error of ± 0·03. In Table II are reproduced the temperature effects obtained by Moon and Tillman, with a cooled wax layer of 1 cm. thickness. Our observations on silver represent the effect for a mixture of the two periods, approximately in equal proportions. The maximum of

J occurs, for all the elements examined, when the thickness of the cooled layer of wax is about 1 cm. A comparison of the results shows that no widely different values of *J* are apparent.

TABLE I

Element	Wax thickness 0·7 cm.	0·9 cm.	1·1 cm.	1·3 cm.
Indium	1·22	1·34	1·42	1·43
Gold	1·26	1·35	1·36	1·33
Sodium	1·42	1·49	1·45	1·35
Arsenic	1·30	1·34	1·39	1·37
Manganese	1·33	1·37	1·44	1·41
Bromine	1·35	1·36	1·53	1·44
Aluminium	1·50	1·54	1·57	1·53
Silver	1·30	1·36	1·40	1·38

TABLE II

Element	*J*
Silver (22 sec.)	1·34 ± 0·03
Silver (140 sec.)	1·35 ± 0·05
Copper (5 min.)	1·47 ± 0·05
Vanadium (225 sec.)	1·47 ± 0·05
Iodine (25 min.)	1·19 ± 0·03
Dysprosium (2·5 hr.)	1·32 ± 0·04
Rhodium	1·25 ± 0·04

Bethe (1935) showed that, for neutrons of energies less than 1000 e-volts, the absorption cross-section of a nucleus was inversely proportional to the neutron velocity. However, experiments on selective absorption have shown the inadequacy of this simple relation; and it has been suggested that the absorption probability also depends upon whether the incident neutron has such an energy that the product nucleus may be formed in one of its excited states (Bohr 1936; Breit and Wigner 1936).

Thus it appears, in general, that the elements which show the strongest selective absorption, show the lower temperature effects, e.g. silver, gold, rhodium, iodine; whilst the elements which are not known to show any pronounced selective absorption, e.g. aluminium, sodium, bromine, tend to show rather higher values of *J*.

It seemed at first as if indium was showing no maximum for *J* as the thickness of the wax layer increased. It was thought that this might be connected with a temperature effect on the neutrons of group D, i.e. the neutrons selectively absorbed by indium. Such an effect (Mitchell, Murphy and Whitaker 1936) has been recorded, and also a similar effect (Rasetti and Fink 1936) in the case of rhodium which also shows selective absorption of group D. In both these cases the temperature range was the same as the one we used, and both gave J for the D neutrons as 1·08.

We therefore performed experiments in which cylindrical cadmium absorbers of thickness 0·3 g./cm.^{2} (i.e. sufficient to absorb all group C neutrons) were placed, one in each of the flasks X and Y with a wax layer 1·1 cm. thick. The cadmium cylinders were closed both at the top and at the bottom, and the indium specimens were suspended inside them. We were unable to detect any temperature effect on the group D neutrons, the value of J obtained under these conditions being $1{\cdot}00 \pm 0{\cdot}01$ (5).

Using the flasks with a wax layer of 1·1 cm. thickness, we determined the fraction of the activity of the specimens remaining after irradiation through a cadmium absorber of thickness sufficient to absorb all group C neutrons (0·3 g./cm.^{2}). These experiments were performed at room temperature. Table III gives the percentage of the total activity in each element, under the conditions of our experiment, due to neutrons other than those of group C.

It is evident that the same number of neutrons of energies greater than thermal will pass through the wax layer whether it be hot or cold. This is supported by our experiments described above in which we found that the activity produced in indium by group D neutrons is independent of temperature.

Assuming that group C consists simply of neutrons of thermal energy we can then calculate the values of J_c for the group C neutrons for each of the elements, by subtracting from each of the counts obtained the number due to non-group C neutrons. For example in the case of indium, for a wax layer of 1·1 cm., $J = 1{\cdot}42 \pm 0{\cdot}01$ (4); and the fraction of the activity at room temperature due to non-group C neutrons is $(17{\cdot}7 \pm 0{\cdot}5)$ %. So that

$$J_c = \frac{142 - 17{\cdot}7}{100 - 17{\cdot}7} = 1{\cdot}52 \pm 0{\cdot}01\ (5).$$

Table III gives the values of J_c, corresponding to a cooled wax layer of 1·1 cm., for the elements investigated.

If the relation between absorption cross-section and neutron velocity were the same for all elements in the thermal region, and if cadmium absorbed only thermal neutrons, then we should expect the J_c values in Table IV to be the same, within experimental error.

The probable errors of the J_c values are given, and several elements appeared to give values of J_c differing considerably from the mean value of 1·51. Manganese and bromine are easily investigated, we therefore repeated the temperature measurements on these two elements in order to reduce the probable error, and so to determine the chances that they showed values of J_c differing from the mean. Table IV shows the final values obtained for these two elements. In the case of bromine the ratio of the deviation from the mean to the probable error of J_c is 5, and for manganese the ratio is 7. This means that the chances that the values of J_c for the two elements are the same as the mean are exceedingly small.

TABLE III

Element	J	% activity not due to group C	J_c
Indium	1·42 ± 0·01 (4)	17·7 ± 0·5	1·52 ± 0·01 (5)
Gold	1·36 ± 0·03	15·6 ± 0·7	1·43 ± 0·03
Sodium	1·45 ± 0·03	6·8 ± 0·7	1·48 ± 0·03
Arsenic	1·39 ± 0·03	22·8 ± 0·6	1·51 ± 0·03
Manganese	1·44 ± 0·02	2·5 ± 0·7	1·45 ± 0·02
Bromine	1·53 ± 0·03	21·1 ± 0·5	1·67 ± 0·03
Aluminium	1·57 ± 0·03	5·2 ± 1·3	1·60 ± 0·03
Silver	1·40 ± 0·03	13·5 ± 0·7	1·46 ± 0·03 (5)

TABLE IV

Element	J	J_c
Manganese	1·41 ± 0·01 (3)	1·42 ± 0·01 (3)
Bromine	1·47 ± 0·01 (6)	1·60 ± 0·01 (8)

These results can be explained in two ways, not mutually exclusive. First, it is possible that the variation of absorption cross-section with velocity for neutrons of thermal velocities is different for different elements. If this be so, most of the elements investigated must show selective absorption for neutrons in the thermal region. Secondly, the absorption law in the thermal region may be the same for all elements, but cadmium may absorb selectively neutrons other than those of thermal energy. In this case all the lower values of J_c would indicate that the elements concerned have an absorption band coinciding more or less with one of those of cadmium; whereas the highest value of J_c obtained indicates that the absorption spectrum of the element in question coincides least with that of cadmium. Consequently this highest value tends most nearly to that to be expected for purely thermal neutrons.

We are grateful to Professor G. P. Thomson for the interest he has taken in the work; and we are indebted to Dr. P. B. Moon, who suggested the problem, for many helpful discussions. We would like also to express our thanks for the gifts of radon from the Medical Research Council, kindly arranged by Professor S. Russ, of the Middlesex Hospital; and to Messrs. G. W. and B. A. Spicer for their skilful preparation of the sources.

SUMMARY

The effect of temperature on the activity produced by thermal neutrons in various elements has been investigated, and no widely differing effects have been observed. The fraction of the activity of each element due to group C neutrons was determined, and the temperature effects due to these neutrons alone were deduced. From the values obtained it appears either that cadmium absorbs neutrons other than those of thermal energy, or that all elements do not conform to the same absorption law in the thermal region. It was also found that the activity produced in indium by group D neutrons was not influenced by temperature.

REFERENCES

Bethe, H. A. 1935 *Phys. Rev.* **47**, 747–59.
Bohr, N. 1936 *Nature, Lond.*, **137**, 344–8.
Breit, G. and Wigner, E. 1936 *Phys. Rev.* **49**, 519–31.
Mitchell, A. C. G., Murphy, E. J. and Whitaker, M. D. 1936 *Phys. Rev.* **50**, 133–7.
Moon, P. B. and Tillman, J. R. 1936 *Proc. Roy. Soc.* A, **153**, 476–92.
Rasetti, F. and Fink, G. A. 1936 *Phys. Rev.* **49**, 642.

(*Abstract*)

The Thermoelastic Properties of Muscle and their Molecular Interpretation

By KURT H. MEYER AND L. E. R. PICKEN

The effect of temperature change on the elastic force exerted by a stretched, unstimulated muscle at constant length has been investigated.

At small and large elongations, the temperature coefficient is negative: the muscle behaves like a normal solid.

At intermediate elongations the coefficient is positive and the elastic force increases more rapidly than the absolute temperature. From this behaviour the following conclusions about the structure of muscle and the origin of the elastic force may be drawn:

(*a*) The elastic system of muscle (and in particular that of the fibrils) consists of flexible, primary-valence chain molecules, the arrangement of which in the resting muscle is random and highly probable, but becomes ordered and less probable in the stretched muscle. Thermal agitation tends to provoke a return to the original, unordered condition.

(*b*) Latent heat is set free during stretching and may be considered as a sort of heat of crystallization.

X-ray studies show that the degree of orientation of the muscle substance is augmented by stretching and diminished by warming—in agreement with the conception of molecular structure reached from a consideration of the thermoelastic behaviour.

A molecular interpretation of the viscous-elastic properties of muscle is offered, and is supported by experimental evidence. The elastic system of the muscle behaves as if composed of two components: flexible protein chains forming a three-dimensional network, and free chains in the meshes of this net.

While in the resting muscle the links of neighbouring chains are able to slide freely over each other, like the molecules in a liquid, in the actively contracting muscle such "liquid" bonds appear to be replaced by more "solid" linkages.

(*The full paper is published in Proc. Roy. Soc. B*, **124**, 29, 1937.)

THE BAKERIAN LECTURE

Regularities and Irregularities in the Ionosphere—I

BY E. V. APPLETON, F.R.S.

Read 10 *June*—*Received* 29 *July* 1937

[Plates 23–28]

Our knowledge concerning the state of the atmosphere lying above about 80 km. in height has been derived from experiments on radio wave reflexion as well as from studies of terrestrial magnetism and of the aurora. The information derived from radio experiments is, fortunately, in the nature of a supplement to, rather than a duplicate of, information derivable in other ways. As one of the best examples in this connexion may be mentioned the question of electrical conductivity. Here the magnetic studies of Schuster and Chapman yield an estimate of the total conductivity for currents travelling horizontally, whereas the radio measurements give the state of ionization at different levels from which the specific conductivity at those levels may be estimated.

One of the most striking things about the ionosphere is the marked solar control. Speaking generally it may be said that the ionization increases and decreases as the sun rises and sets. Again, speaking generally, we may say that the main part of the ionization is caused by solar ultra-violet light. The rays from the sun meet the outer layers of the atmosphere first and the short wave-length radiation is absorbed there, causing ionization. It thus comes about that the study of the ionosphere becomes the study of an interesting part of the sun's spectrum which cannot be detected at ground level. It also becomes the study of certain atomic processes such as photo-ionization, recombination of ions and attachment of electrons to neutral molecules such as cannot be investigated at very low pressure in the laboratory, because of the influence of the walls of the vessel confining the gas.

One further word may be added by way of introduction. From the way in which the air molecular density falls with increase of height in the lower atmosphere we can estimate, assuming a continuance of the exponential decay with height, what the molecular density should be at a height of 300 km. The value found (of the order of 1000 mol./c.c.) is considerably less than the ionization densities found at such a height (of the order of 10^6 electrons per c.c.). A substantial part of the ionosphere must therefore be situated in regions where the proportional fall of pressure with increase of height is considerably smaller than it is at lower levels. In other words there must be some influence at work tending to spread the atmosphere out to great heights. The nature of this influence is considered in greater detail below.

In the present communication selected topics from the information derived from radio sounding of the ionosphere are discussed and considered in connexion with results derived from a theory of simple layer formation by solar ionizing radiation, travelling rectilinearly and attenuated according to a mass-absorption law. The comparisons are concerned chiefly with the variation of maximum electron content N_m with solar radiation angle of incidence χ, with the total electric conductivity for direct currents such as determines the quiet day magnetic variations and with the high-frequency absorption of radio waves traversing such a simple layer, the last two quantities being, to some extent, related.

Other matters discussed are (*a*) ionization decay processes in the ionosphere, (*b*) the actual height at which radio waves are reflected and (*c*) ionospheric disturbances.

1—SIMPLE REGION FORMATION

The elucidation of ionospheric phenomena by radio methods has been greatly assisted by the comparison of experimental results with those to be expected from upper-atmospheric ionization by homogeneous solar radiation. The essential formula relating ion-production to the angle of incidence of the solar radiation was first given by Lenard (1911), who showed that in an ideal atmosphere the ionization by a radiation suffering mass absorption always led to layer formation with a lower boundary sharper than the upper boundary.

In two papers of fundamental importance to ionospheric studies, Chapman (1931 *a*, *b*) has given a complete theory of layer formation so as to include the effects of diurnal and seasonal variations of the angle of incidence of the solar rays as well as the influence of recombination. For our present purpose, however, only the result of Lenard's calculations is required, although the convenient notation introduced by Chapman is employed.

We measure all atmospheric heights (h) as differences, positive and negative, from a certain datum level h_0 which is the height at which noon ion-production is a maximum at the equator at the equinox. It is also convenient, still following Chapman, to measure these height-differences (z) in terms of H, the scale height, for the region in question. We then have

$$z = \frac{h - h_0}{H}, \tag{1}$$

where

$$H = \frac{KT}{mg}, \tag{2}$$

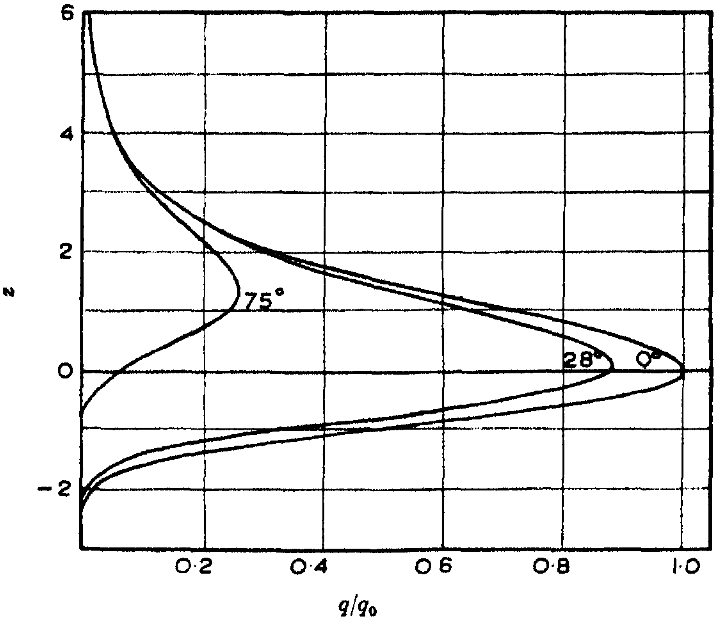


FIG. 1—Curves showing the relation between ion-production and height for south-east England in summer ($\chi = 75°$) and winter ($\chi = 28°$), and for the equator at equinox noon ($\chi = 0$).

K, m and g denoting respectively Boltzmann's constant, the mean molecular mass and the acceleration due to gravity, while T is the absolute temperature. The molecular density, n, and the electron collision frequency, ν, will correspondingly be $n_0 e^{-z}$ and $\nu_0 e^{-z}$ respectively, where n_0 and ν_0 are the relevant values at the datum level h_0. It is here assumed, for simplicity, that the scale height for the component ionized is the same as that for the gas as a whole.

The relation between rate of ion-production q, height z and angle of incidence χ of the ionizing radiation is

$$q = q_0 e^{1-z-e^{-z}\sec\chi}, \tag{3}$$

where q_0 is the ion-production at height h_0 when $\chi = 0$. In fig. 1 are shown plots of q for the three cases $\chi = 0$ (equator at equinox noon), $\chi = 28°$ (south-east England at summer noon) and $\chi = 75°$ (south-east England at winter noon). Certain characteristics of such layers of ion-production will now be examined.

(a) *Height of Maximum of Ion-production*

By simple differentiation it is seen that q is a maximum when $\log \cos \chi = -z$. This means that in summer in south-east England z, for maximum q, is 0·125 and in winter, 1·352. Thus the height of maximum ion-production increases by $1{\cdot}23\,H$ from summer noon to winter noon.

The molecular density at the level of maximum ion-production varies inversely with H (and thus inversely with T) and also inversely with the absorption coefficient A of the ionizing radiation. The production of more than one layer in a given atmosphere may therefore be due to the presence of atmospheric constituents of different ionization potentials and thus different A values.* The heights at which different layers are formed are also determined by the values of H which depend on temperature and molecular mass. It is in terms of such quantities that an explanation of the heterogeneous structure of the ionosphere must be sought.

(b) *Height of Under-boundary of Layer of Ion-production*

From a second differentiation of (3) we can find the height of the point of inflexion on the lower part of the curve. This is given by the relation

$$e^{-z} \sec \chi = \frac{3+\sqrt{5}}{2} = 2{\cdot}62.$$

For the special cases of summer and winter noon in England we have $z = -0{\cdot}84$ and $+0{\cdot}39$ respectively, so that we may say that the lower boundary of the layer is $1{\cdot}23\,H$ higher at summer noon than at winter noon.

(c) *Variation of Ionization with Height*

If we assume a recombination coefficient α, constant throughout the layer, we have, at noon or at any time when quasi-equilibrium conditions are established,

$$q = \alpha N^2,$$

---

* Although radiation of a definite absorption coefficient cannot form more than one layer, it is easily seen that a quasi-dual structure is produced if such radiation is absorbed in a region where the value of H changes sharply.

where N is the electron concentration, from which, embodying (3) follows

$$N = \sqrt{\frac{q_0}{\alpha}}\, e^{\frac{1-z-e^{-z}\sec\chi}{2}} \tag{4}$$

or

$$N = N_0 e^{\frac{1-z-e^{-z}\sec\chi}{2}}, \tag{5}$$

where N_0 is the electron concentration at the maximum of ion-production (or of ionization) when $\chi = 0$. A region in which the distribution of ionization with height and with solar angle of incidence is described by (5) may be called a simple region.

(*d*) *Total Ion-production in a Simple Region*

From (3) we see that the total number of ion pairs produced per sec. per cm. column of the atmosphere (Q) is given by

$$\begin{aligned} Q &= q_0 H \int_{-\infty}^{+\infty} e^{1-z-e^{-z}\sec\chi} . dz \\ &= 2{\cdot}72\, q_0 H \cos\chi. \end{aligned} \tag{6}$$

The magnitude of this quantity is of importance when it is desired to compare it with the number of available ionizing quanta in the complete radiation from the sun (cf. Saha 1937).

(*e*) *Total Transverse Conductivity of a Simple Region*

The horizontal specific conductivity σ for steady electromotive forces acting at right angles to the meridian is given by

$$\sigma = \frac{Ne^2}{m} . \frac{\nu}{\nu^2 + p_H^2}, \tag{7}$$

where the medium has N charges of mass m and charge e per cm.^{3}, ν is the collisional frequency and p_H the angular gyro-frequency of the charges round the lines of magnetic force. The charges may be of electronic or molecular mass. If both electrons and ions are present the conductivities are additive. For a simple region the total conductivity, transverse to the magnetic field, is given by

$$\int \sigma\, dh = \frac{N_0 e^2 H \nu_0}{m} \int_{-\infty}^{+\infty} \frac{e^{\frac{1-3z-e^{-z}\sec\chi}{2}}}{\nu_0^2 e^{-2z} + p_H^2}\, dz. \tag{8}$$

After some rearrangement this may be written*

$$\int \sigma\, dh = \frac{N_0 e^2 H}{m}\left(\frac{e^1}{\nu_0 p_H}\right)^{\frac{1}{2}} X, \tag{9}$$

where
$$X = \int_0^\infty \frac{t^{\frac{1}{2}}}{1+t^2} e^{-at}\, dt \tag{10}$$

and
$$a = \frac{p_H \sec\chi}{2\nu_0}. \tag{11}$$

When $a \ll 1$,

$$X = \frac{\pi}{\sqrt{2}}(\cos a + \sin a) - 2\sqrt{\pi a}(1 - \tfrac{4}{15}a^2) + O(a^4). \tag{12}$$

On the other hand, when $a \gg 1$,

$$X = \tfrac{1}{2}\sqrt{\frac{\pi}{a^3}}, \tag{13}$$

so that
$$\int \sigma\, dh = \sqrt{2\pi e^1}\,\frac{e^2}{m}\,\frac{N_0 \nu_0 H}{p_H^2}\cos^{\frac{3}{2}}\chi. \tag{13a}$$

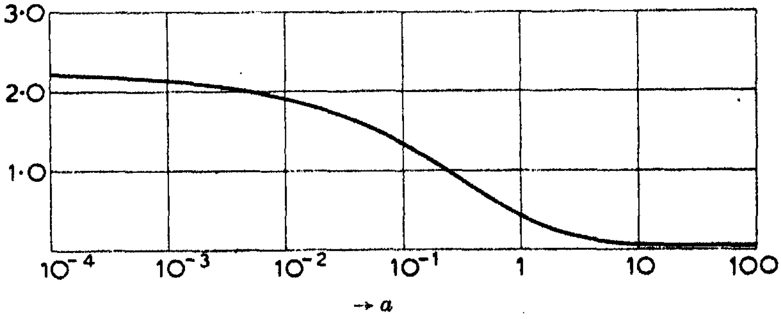


FIG. 2

For the evaluation of X when a is of the order of unity recourse must be had to integration by quadrature.† The graph relating X and a for values intermediate between the ranges of validity of (12) and (13) is shown in fig. 2. For our present purpose we are interested in the determination of the conductivity of simple regions of electrons or of ions situated at different heights (i.e. for different values of ν_0). For simplicity we take the case of $\chi = 0$.

For electrons $p_H = 8{\cdot}3 \times 10^6$ c./sec., so that the layer conductivity may

---

[*] I am much indebted to Mr. M. V. Wilkes for assistance in this connexion.

[†] For these evaluations I am indebted to Dr. L. J. Comrie's Scientific Computing Service.

be calculated with the aid of (12), (13) and the data in fig. 2 for different values of ν_0, the collisional frequency at the height of maximum ionization. For ions of average mass 28 the corresponding value of p_H is 160·4 c./sec. The results of such calculations are given in Table I.

TABLE I

ν_0	Region conductivity (electrons) e.s.u.	Region conductivity (ions) e.s.u.
10^1	$1{\cdot}508 \times 10^{-4}\ N_0H$	$7{\cdot}29\ N_0H$
10^2	$1{\cdot}508 \times 10^{-3}\ N_0H$	$39{\cdot}5\ N_0H$
10^3	$1{\cdot}508 \times 10^{-2}\ N_0H$	$28{\cdot}12\ N_0H$
10^4	$1{\cdot}508 \times 10^{-1}\ N_0H$	$12{\cdot}36\ N_0H$
10^5	$1{\cdot}508\ N_0H$	$4{\cdot}30\ N_0H$
10^6	$11{\cdot}51\ N_0H$	$1{\cdot}40\ N_0H$
10^7	$32{\cdot}98\ N_0H$	$0{\cdot}448\ N_0H$
10^8	$22{\cdot}88\ N_0H$	$0{\cdot}142\ N_0H$

(*f*) *Absorption of Radio Waves Traversing a Simple Region*

From the point of view of experiments to be described later we require to know the absorption suffered by waves travelling through a simple region. In this case an electron region only need be considered, since the high frequency absorption of an ion layer is negligible in comparison. A special case of this arises in the consideration of the attenuation of waves reflected by Region F in their journey up and down through Region E.

Now it has previously been shown that the main absorption suffered by such waves is due to the lower part of Region E, where the collisional frequency is relatively high and yet the refractive index of the medium not substantially different from unity. In such a case we may write for the spatial absorption coefficient κ, for vertical propagation,

$$\kappa = \frac{2\pi e^2}{mc}\left(\frac{N\nu}{\nu^2 + (p \pm |p_L|)^2}\right), \tag{14}$$

where $|p_L|$ is the electron angular gyro-frequency corresponding to the vertical component of the earth's magnetic field ($7{\cdot}91 \times 10^6$ c./sec. in England), and the upper and lower signs refer to the ordinary and extraordinary components respectively. For simplicity we shall consider only the case of the former.

For a simple region (e.g. day-time, Region E) we may include the variation of N and ν with height and (14) becomes

$$\kappa = \frac{2\pi e^2}{mc} N_0 \nu_0 \left\{\frac{e^{\frac{1-3z-e^{-z}\sec\chi}{2}}}{\nu_0^2 e^{-2z} + (p + |p_L|)^2}\right\}, \tag{15}$$

so that if we regard the diminution below unity of the reflexion coefficient ρ of, say, Region F, as due to the double traverse of Region E, we may write

$$-\log\rho = \frac{4\pi e^2 N_0 \nu_0 H}{mc}\int_{-\infty}^{+\infty} \frac{e^{\frac{1-3z-e^{-z}\sec\chi}{2}}}{\nu_0^2 e^{-2z}+(p+|p_L|)^2}.dz \tag{16}$$

$$= \frac{4\pi e^2 N_0 H}{mc}\left(\frac{e^1}{\nu_0(p+|p_L|)}\right)^{\frac{1}{2}} X, \tag{17}$$

where X is given by (10) and $a = \dfrac{(p+|p_L|)\sec\chi}{2\nu_0}$. Since X is known in terms of a, the value of ρ can be calculated for any simple electron region situated at any height.

We proceed to consider what is probably a sufficiently good approximation for the case of Region E absorption of waves of frequency sufficiently high to penetrate it. In this case we assume that the collisional frequency is small compared with the electric wave frequency, so that a is large. For such conditions, (15) and (16) become respectively

$$\kappa = \frac{2\pi e^2 N_0 \nu_0}{mc(p+|p_L|)^2} e^{\frac{1-3z-e^{-z}\sec\chi}{2}} \tag{18}$$

and

$$-\log\rho = 4{\cdot}13\left(\frac{4\pi e^2}{mc}\right)\frac{N_0 \nu_0 H \cos^{\frac{3}{2}}\chi}{(p+|p_L|)^2}. \tag{19}$$

From (18) it is easily shown that κ, the absorption factor, is a maximum when e^{-z} is equal to $3\cos\chi$. There is thus a simple relation between the height of maximum ionization and height of maximum absorption. Thus let z_κ be height of maximum absorption and z_N be height of maximum ionization. We have

$$e^{-z_\kappa} = 3\cos\chi \quad \text{and} \quad e^{-z_N} = \cos\chi,$$

so that

$$e^{(z_N - z_\kappa)} = 3. \tag{20}$$

Thus the level of maximum absorption is $1{\cdot}1\,H$ below the level of maximum ionization.

Equation (19) may be rewritten in terms of the ordinary ray critical frequency f_E for Region E as follows:

$$-\log\rho = \frac{4{\cdot}13}{c}\left(\frac{f_E}{f+|f_L|}\right)^2 H\nu_0\cos\chi, \tag{21}$$

where f is the electric wave frequency, $|f_L|$ the longitudinal gyro-frequency and c the velocity of light.

Finally an interesting relation between the direct current conductivity and the high-frequency absorption of a simple electron region may be noted. From (13*a*) and (19) it is easily shown that

$$\int \sigma\, dh = \frac{c}{4\pi}\left(\frac{p+|p_L|}{|p_H|}\right)^2 |\log \rho|. \qquad (22)$$

Thus the total conductivity of a simple electron region may be estimated from the reflexion coefficient of waves traversing that region and reflected from a higher region. The limitation of this formula to a case where the collision frequency is small compared with the electric wave frequency should be borne in mind.

2—EXPERIMENTAL RESULTS AND DISCUSSION

(*a*) *The Variation of Ionization with the Sun's Zenith Distance*

Region E of the ionosphere, in the middle of the day, is a good example of a simple region as defined above. Here the maximum electron content is throughout the day maintained close to the equilibrium value $\sqrt{q/\alpha}$.

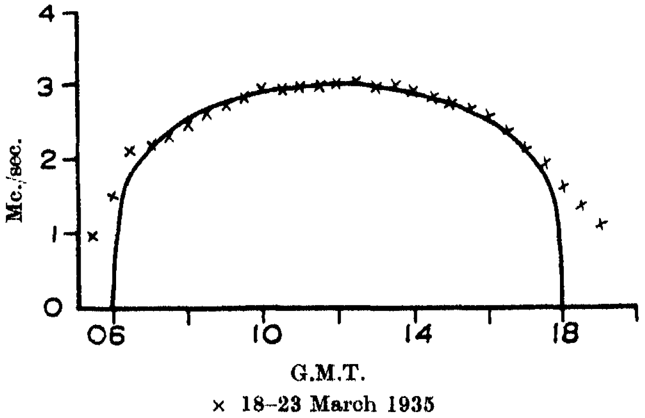


FIG. 3—Diurnal variation of average Region E critical frequency for the period 18–23 March 1935; continuous curve represents variation of $(\cos\chi)^{\frac{1}{4}}$.

Under such conditions the value of $N_{\max}$ should, according to (5), vary as $(\cos\chi)^{\frac{1}{2}}$, so that the critical penetration frequency which is proportional to $(N_{\max})^{\frac{1}{2}}$ should vary as $(\cos\chi)^{\frac{1}{4}}$. How closely this is the case is illustrated in fig. 3, where the experimental average values of f_E are plotted over the day-time hours for the period 18–23 March 1935. In the same graph is shown as a continuous line the variation of $(\cos\chi)^{\frac{1}{2}}$, identity of ordinates being adjusted at mid-day.

Now the maintenance of N at the equilibrium values indicates that dN/dt is small compared with q or αN^2, t being the time. We can therefore estimate an inferior limit for α. Now during the period 1600 to 1700, the value of dN/dt is 7·4 electrons per c.c. per sec. while at 1630 the value of αN^2 is $4{\cdot}9 \times 10^9\alpha$. Thus in order that dN/dt should be less than, say, a tenth of αN^2, the value of α must be greater than 10^{-8}. This is greater than the values of α (of the order of 2×10^{-9}) found for night-time conditions. There appears therefore to be a marked increase in the value of the effective electron recombination coefficient from night to day.

Region F_1 behaves similarly to Region E and so can, during the day, be classed as a simple region. On the other hand, in Region F_2 the value of α is sufficiently small during a summer or equinox day to cause the ionization to depart from the equilibrium value. During a winter day Region F_2 behaves more nearly as a simple region.

(b) *The Variation of Absorption with the Sun's Zenith Distance*

Much attention has been paid in recent years to the systematic measurement of equivalent heights of reflexions and ionospheric electron densities, though corresponding measurements of the reflexion coefficients for the various reflecting strata are relatively few. This is rather remarkable in view of the great practical interest of the latter to the radio-engineer.

The measurements to be described are relatively simple and any value they possess arises from their sustained nature. They were carried out in order to test how the absorption of radio waves varied with χ over the seasons, and consist of measurements of the reflexion coefficient of Region F for 4 Mc./sec. waves. This frequency was chosen because it was higher than the values of Region E critical frequency then experienced (1934) and yet sufficiently low not to approach the Region F penetration value. All observations were made at noon.

The sending station was accommodated in the wireless hut on the roof of the east wing of King's College, London, while the receiving apparatus was installed at the Halley-Stewart Laboratory, Hampstead. In the course of each run of measurements twenty snapshots of the echo-pattern were made at half-minute intervals. From these the mean ratio of first echo to ground amplitudes, F'/G, was easily found, since the receiver, tested by means of a signal generator, had a sensibly linear response of rectified output to high-frequency input. After some months of preliminary runs in 1934 it was considered that the results obtained justified a series lasting at least two or three years. The measurements were therefore continued throughout 1935 and 1936, for the first 21 months of this period in London and for the last three at Cambridge.

In the London series the absorption of the ground waves from the Strand to Hampstead was sufficient to make the intensities of the G and F' (and F'' when observed) amplitudes of comparable magnitude when the amplifier gain was kept at a constant value. Thus the ratio of F'/G could be found for each photographic record by simple measurement. This quantity was therefore proportional to the reflexion coefficient ρ. To find the constant of proportionality similar observations were made in the evenings when the F' echo was greater in amplitude and the amplitude of the second reflexion F'' was appreciable. Allowing for the fact that the doubly reflected wave has made a double journey to the reflecting region and back the value of ρ is then given by $2(F''/F')$, if no allowance is made for loss of amplitude at ground reflexion, or $2F''/\rho'F'$ if ρ' is the ground amplitude reflexion coefficient for the wave frequency employed. The constant β in the relation $\rho = \frac{\beta F'}{G} = \frac{2F''}{\rho' F'}$ could then be determined from a large number of simultaneous measurements of F'/G and F''/F'.

At Cambridge the conditions were such that correspondingly simple methods of calibration could not be employed. Here the sender was situated on the Solar Physics Observatory site only 1·28 km. from the receiving station which was set up in the wireless hut on the Cambridge University Rifle Range. Under these conditions the ground ray amplitude was so large that the amplitudes of G and F' were not usually comparable for a constant setting of the amplifier gain. The following method was therefore employed. The gain control of the intermediate frequency amplifier used was calibrated by means of a signal generator, the gain relative to the zero reading being expressed in nepers.

In operation the gain of the amplifier was adjusted to bring, in turn, the amplitudes of G and F' (and F'', F''', etc. where observable) to a specified mark on the cathode-ray oscillograph. From the gain readings, the values of $\log(F'/G')$, $\log(F''/F')$, $\log(F'''/F'')$, etc. could then be read off the calibration curve.

Now since the reflexion coefficient ρ is given by

$$\rho = \frac{\beta F'}{G} = \frac{2F''}{\rho' F'},$$

we have

$$\log \rho = \log \beta + \log \frac{F'}{G} = \log 2 - \log \rho' + \log \frac{F''}{F'}.$$

From simultaneous measurements of $\log(F'/G)$ and $\log(F''/F')$ the value of $\log\beta$ was found so that $\log\rho$ could be estimated when first reflexions alone were present. It will be seen that the value of $\log\beta$ depends on the nature and relative dispositions of the sending and receiving aerials. A new determination of its value was always made when, for other reasons, changes were made in these installations. It will be seen that a great advantage of the second method of calibration is that the reflexion coefficient is expressed logarithmically in nepers, as is required for comparison with theory.

In fig. 4 is shown the curve of monthly averages of ρ for the years 1935 and 1936.

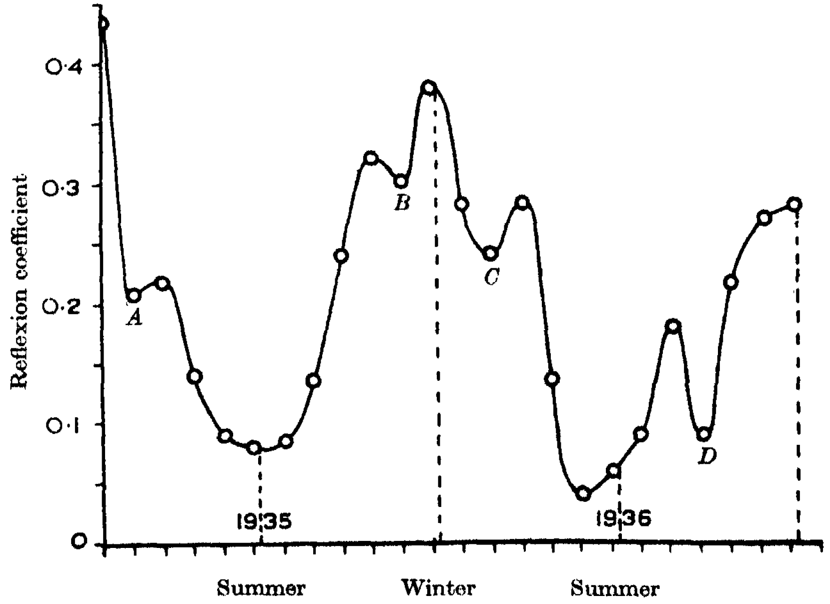


FIG. 4—Monthly means of ionospheric reflexion coefficient (ρ) for 4 Mc./sec. waves.

The graph shows very clearly the greater absorption in summer than in winter, the average value of $\log\rho$ being about -1 neper in winter and -3 nepers in summer. It is clear from the graph that the results are explicable in a general way in terms of an absorbing layer which is more densely ionized in summer than in winter.

The very marked absorption indicated at points A, B, C and D is due to the influence of Region F_1, for at these times the frequency of 4 Mc./sec. was the critical value for this region. In such cases, as has been shown by Eckersley (1935), Farmer and Ratcliffe (1935), and also by White and Brown (1936), there is an additional effect of absorption due to group retardation.

The ratio of the summer and winter values of $\log \rho$, found in these experiments to be about 2·6, is an important ionospheric quantity, and it is of interest to examine the long-distance reception data to see if any confirmation of its order of magnitude can be obtained.*

There is, of course, no lack of evidence that short-wave signals are stronger in winter than in summer in conditions under which electron-limitation is not operative. Quantitative results are given, in this connexion, by Burrows (1931), in a series of graphs illustrating reception data for the propagation of short radio waves over the North Atlantic. Burrows's results show that, during the day-time, as the frequency of the wave is increased the signal strength increases (when expressed in field strength per kilowatt radiated) until a maximum value is obtained, after which the signal falls sharply.

Expressing the attenuation as $e^{-\gamma}$ it may be shown that, for the lower frequencies, γ is approximately inversely proportional to the square of the frequency, which would, according to the theory given above, suggest absorption in a layer in which $\nu^2 \ll p^2$. According to the views expressed above this region lies below the deviating portion of Region E.

To find a ratio of summer to winter attenuation let us express the attenuation coefficient as $e^{-\delta/f^2}$. If now we find, from Burrows's results, summer and winter frequencies which give the same attenuation (i.e. the same signal strength per kilowatt radiated), we have

$$\frac{\delta_s}{\delta_\omega} = \frac{f_s^2}{f_\omega^2}.$$

Using Burrows's measurements for the years 1926–28 we can show from a number of such comparisons that the value of δ_s/δ_ω ranges from 2·1 to 2·7, with an average of 2·4. It will be seen that this figure is the same as the ratio of the summer and winter values of $\log \rho$, suggesting that the two cases of vertical and oblique incidence are affected by the same absorbing region.

Now, as has been shown above, the absorption of radio waves in traversing a particular region is proportional to the direct-current conductivity for horizontal electromotive forces. It is also known that the quiet day magnetic variation is two or three times as large in summer as in winter. The absorption results therefore suggest that the magnetic currents flow in the lower part of the ionosphere.

The theory given in (1f) indicates the general relation between $|\log \rho|$ and χ. For the special case where $\nu^2 \ll (p + |p_L|)^2$ it is shown that $|\log \rho|$

---

* I am indebted to Mr. T. L. Eckersley who has communicated to me privately that he finds a similar average value for the summer to winter ratio of $\log \rho$ for frequencies between 3 and 5 Mc./sec.

varies as $(\cos\chi)^{\frac{3}{2}}$. In fig. 5 is illustrated a test of this relation for 1935, identity of ordinates being arranged at midsummer. The corresponding variation of $\cos\chi$ is also shown for comparison. Although fair agreement is obtained for the summer months it will be seen that the experimental and theoretical curves diverge for the winter months. This discrepancy can also be seen if the experimental results are viewed in another way. For south-east England the ratio of the summer and winter values of $(\cos\chi)^{\frac{3}{2}}$ is 6·4, whereas, as we have seen, the ratio of the summer and winter values of $|\log\rho|$ is 2·6. There are two possible explanations of this discrepancy. In the first place it is possible that group retardation effects due to Region F_1 are still of appreciable magnitude at mid-winter. In the second place it can easily be shown from (17) that the presence of ionization at a level where ν is of the same order of magnitude as $(p + |p_L|)$ would cause absorption which would not vary with $\cos\chi$ as rapidly as does the term $(\cos\chi)^{\frac{3}{2}}$. Further work is probably required to decide between these two alternatives, though in favour of the former it may be mentioned that Best and Ratcliffe, working at Cambridge, have found good agreement with the relation (19), so far as the relation between $|\log\rho|$ and $\cos\chi$ is concerned, for the case of the diurnal variation on ionospherically undisturbed days.

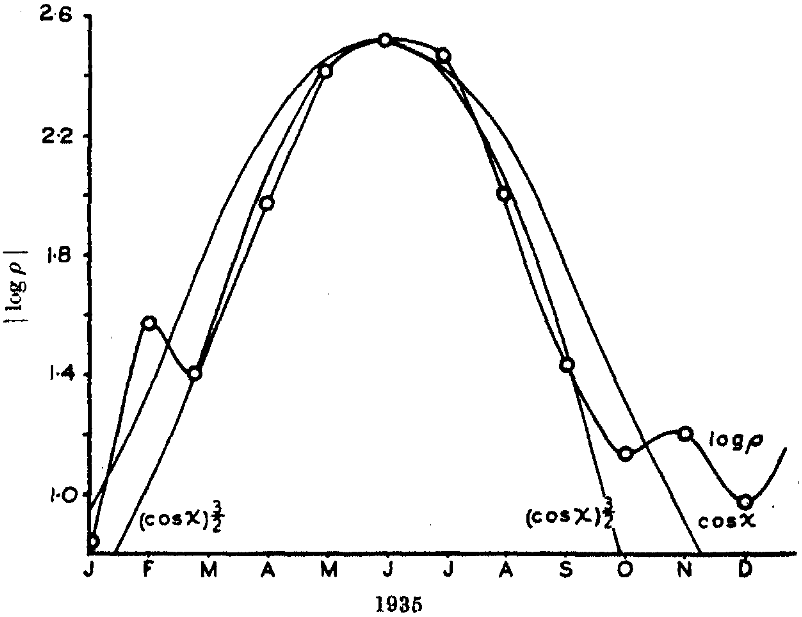


Fig. 5—Values of $|\log\rho|$ for the summer months of 1935 with plots of $\cos^{\frac{3}{2}}\chi$ and $\cos\chi$ for comparison.

(c) *The Direct-current Conductivity of Region E*

As has been suggested by the absorption experiments we consider Region E (and possibly, to some extent, Region F_1) as the location of the currents causing the magnetic variations. Additional evidence in this connexion is the fact that the ionization in Regions E and F_1 varies by about 50 % over the sunspot cycle, as does the amplitude of the magnetic variation. The variation of Region F_2 ionization, on the other hand, exhibits no parallelism with this amplitude, either in respect of seasonal or sunspot cycle changes, indicating that the magnetic currents do not flow at high levels. In south-east England the maximum ionization in Region F_2 has been found to be greater in December than in June. Also the variation over the sunspot cycle is very much greater than 50 %.

It is of interest to see whether the electron conductivity of Region E for steady electromotive forces is of the same order (9×10^{15} e.s.u.) as has been deduced by Schuster and Chapman from their studies of terrestrial magnetism. We take the case of a summer day when $|\log \rho|$ for 4 Mc./sec. waves is of the order of 3. Substituting in (22), where p_H is 8·4 Mc./sec. and where $(p + |p_L|)$ is 33 Mc./sec., we find that $\int \sigma \, dh$ is $1{\cdot}2 \times 10^{11}$ e.s.u. Thus the total electron conductivity is about 8×10^4 times too small to account for the magnetic variations. It is true that Pekeris (1937) has recently shown that the amplitudes of the velocity of the solar semi-diurnal tides at the 100 km. level has probably been previously underestimated by a factor of up to 200, in which case the magnetic data require a conductivity of $4{\cdot}5 \times 10^{13}$ instead of the figure 9×10^{15} e.s.u. But a considerable discrepancy yet remains.

Another way of estimating the order of magnitude of the total electron conductivity is to use Table I. Taking N_0 as $1{\cdot}5 \times 10^5$ electrons per c.c. and H as 10^6 cm., we find that electron regions with ν_0 equal to 10^4 or 10^5 (the appropriate order of magnitude) have total conductivities of $2{\cdot}25 \times 10^{10}$ or $2{\cdot}25 \times 10^{11}$. These values, like the estimate from the high-frequency absorption, are too small. The contributions due to an equivalent number of ions would, taking ν_0 now as 10^3 or 10^4, be $4{\cdot}2 \times 10^{12}$ and $1{\cdot}86 \times 10^{12}$, which are still too small even if the correction suggested by Pekeris is valid. A way out of the difficulty is to assume that there are in Region E, undetected by radio methods, considerable quantities of negative ions which contribute substantially to the direct-current conductivity. The ratio of the number of ions to the number of electrons required will be from 10 to 10^3 according to the magnitude of the Pekeris correction.

It will be seen from Table I that the electron conductivity in Region F_2

must still be small in spite of the fact that N_0 and H (see p. 471) are greater than in the case of Region E. The contribution to the conductivity for an equivalent number of ions is considerably greater. Thus for ν_0 values of 10^2 and 10^3 the total ion conductivity will be (taking N_0 as 10^6 and $H = 50$ km.) $2{\cdot}0 \times 10^{14}$ and $1{\cdot}4 \times 10^{14}$ e.s.u. respectively, which values are intermediate between the magnetic value with and without the Pekeris correction. The fact that Region F_2 is not the location of the magnetic currents indicates therefore that there cannot be a marked excess of ions in Region F_2, and the evidence as a whole suggests that there is little or no negative ion formation there.

(d) *The Process of Electron Production and Decay*

Ionization measurements, made during solar eclipses, have given ample proof that the Regions E and F_1 of the ionosphere are caused by photo-ionization. This may be the liberation of electrons from neutral atoms and molecules or from negative ions. For the case of Region F_2, however, the interpretations of the results obtained are conflicting. Most investigators conclude that it is also caused by ultra-violet light, but Berkner has adduced arguments for neutral particles as the ionizing agency.

Concerning the process of electron disappearance in the ionosphere there is also a conflict of opinion. In a previous paper Appleton and Naismith (1935) have shown that the seasonal variation of the noon ionization for Regions E and F_1 observed by them was most suitably explained in terms of a theory, according to which the electrons disappeared by recombination. This view was challenged by Martyn and Pulley (1936) who contended that the electrons became ineffective by their attachment to neutral atoms, their mobilities, and thus their efficiency in radio wave reflexion, being thereby reduced. Put briefly, the suggestion of Martyn and Pulley was that attachment takes place to oxygen atoms but not to oxygen molecules and that the supply of oxygen atoms is maintained in a sufficiently low concentration by the process of association to form oxygen molecules.

For night-time conditions Martyn and Pulley adopt an entirely different theory. They realize that the attachment process is so rapid that if it occurred alone without some compensating influence, the electrons would disappear in a very short time, whereas, as has been shown, the decay is relatively slow. They therefore make the suggestion that the detachment of electrons must take place as a result of collisions between negative oxygen ions and neutral oxygen atoms, oxygen molecules and free electrons being thereby produced. This difficulty of having to adopt different theories for the day and night conditions is a very real one, since the day-time

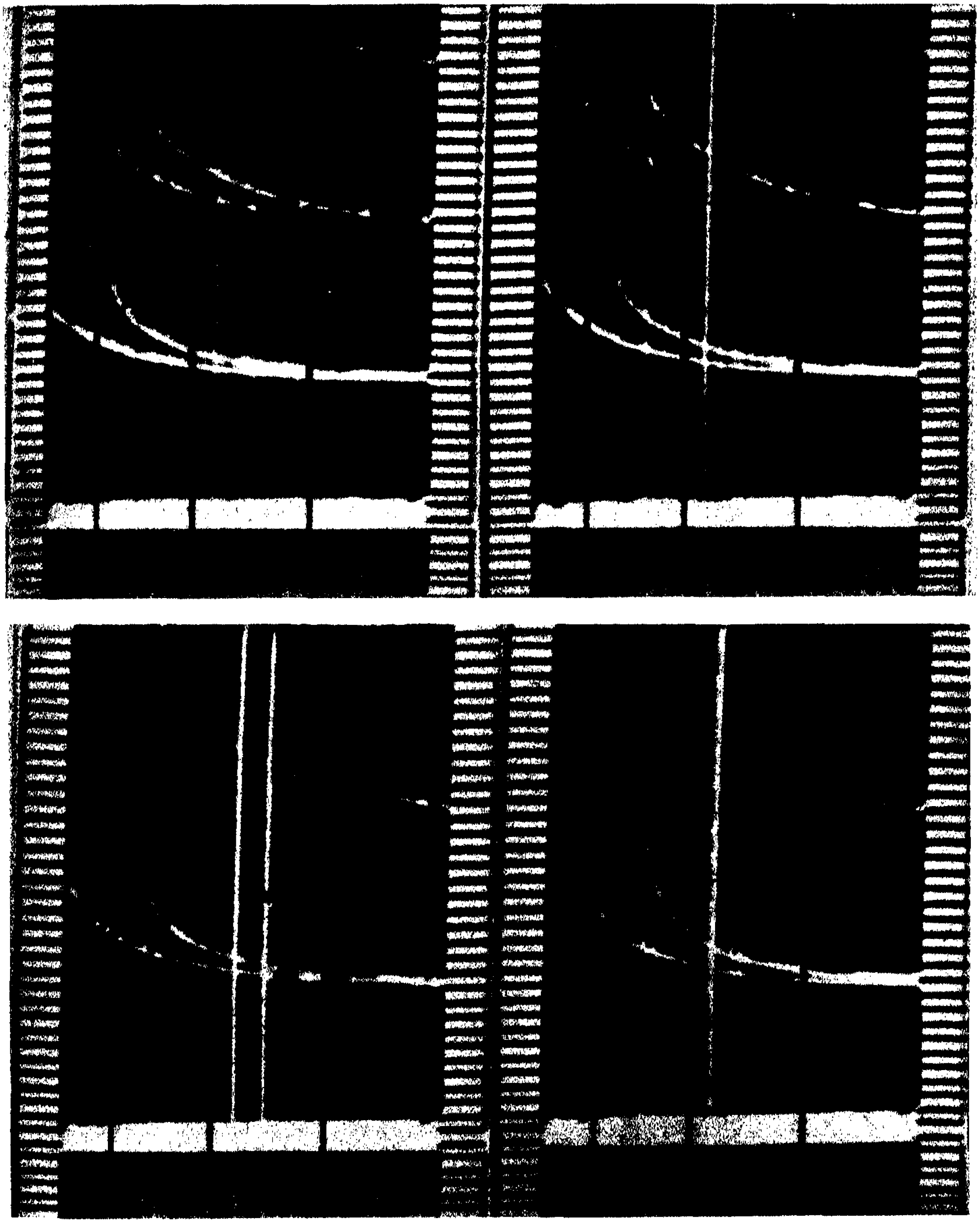


Fig. 6

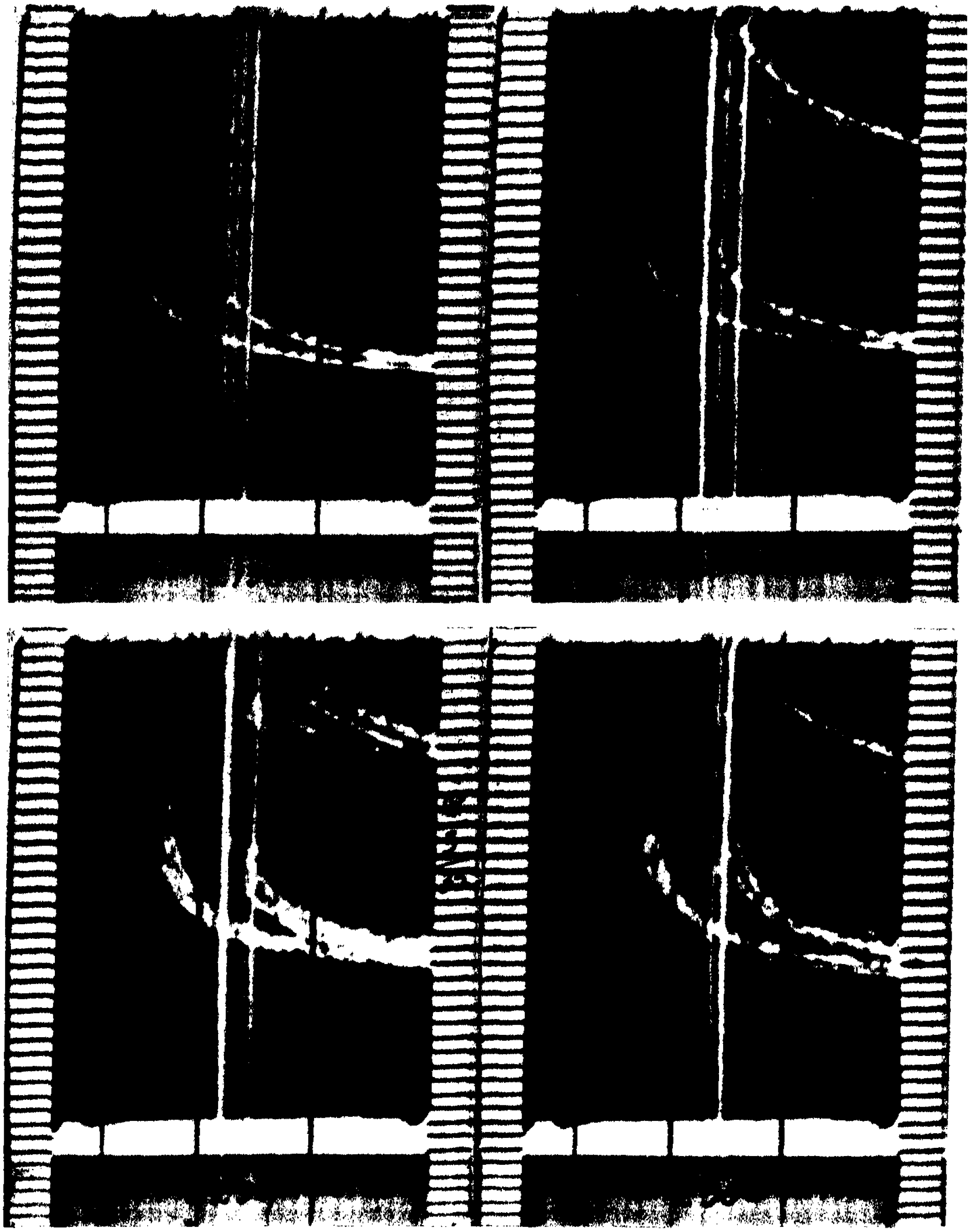


FIG. 6

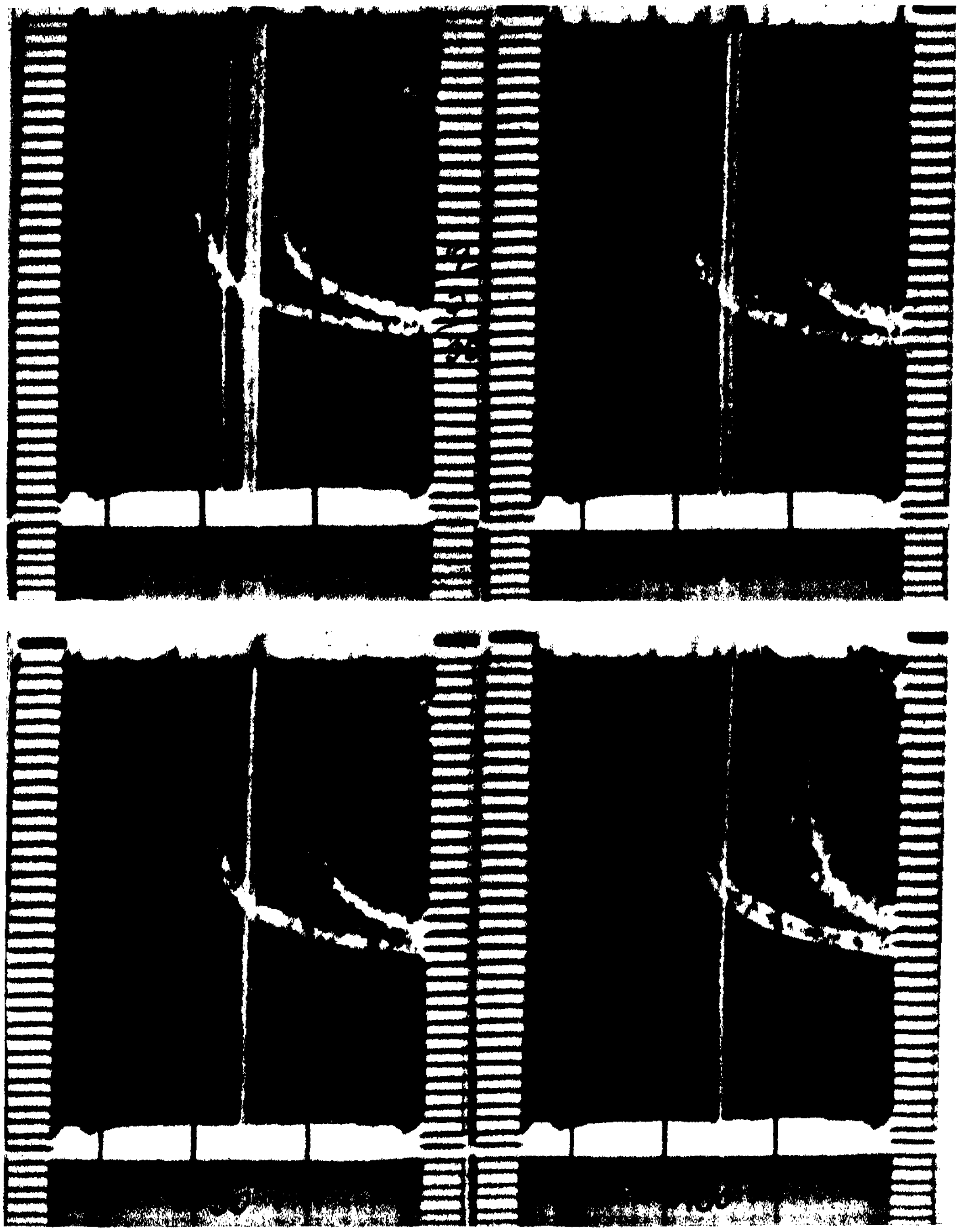


FIG. 6

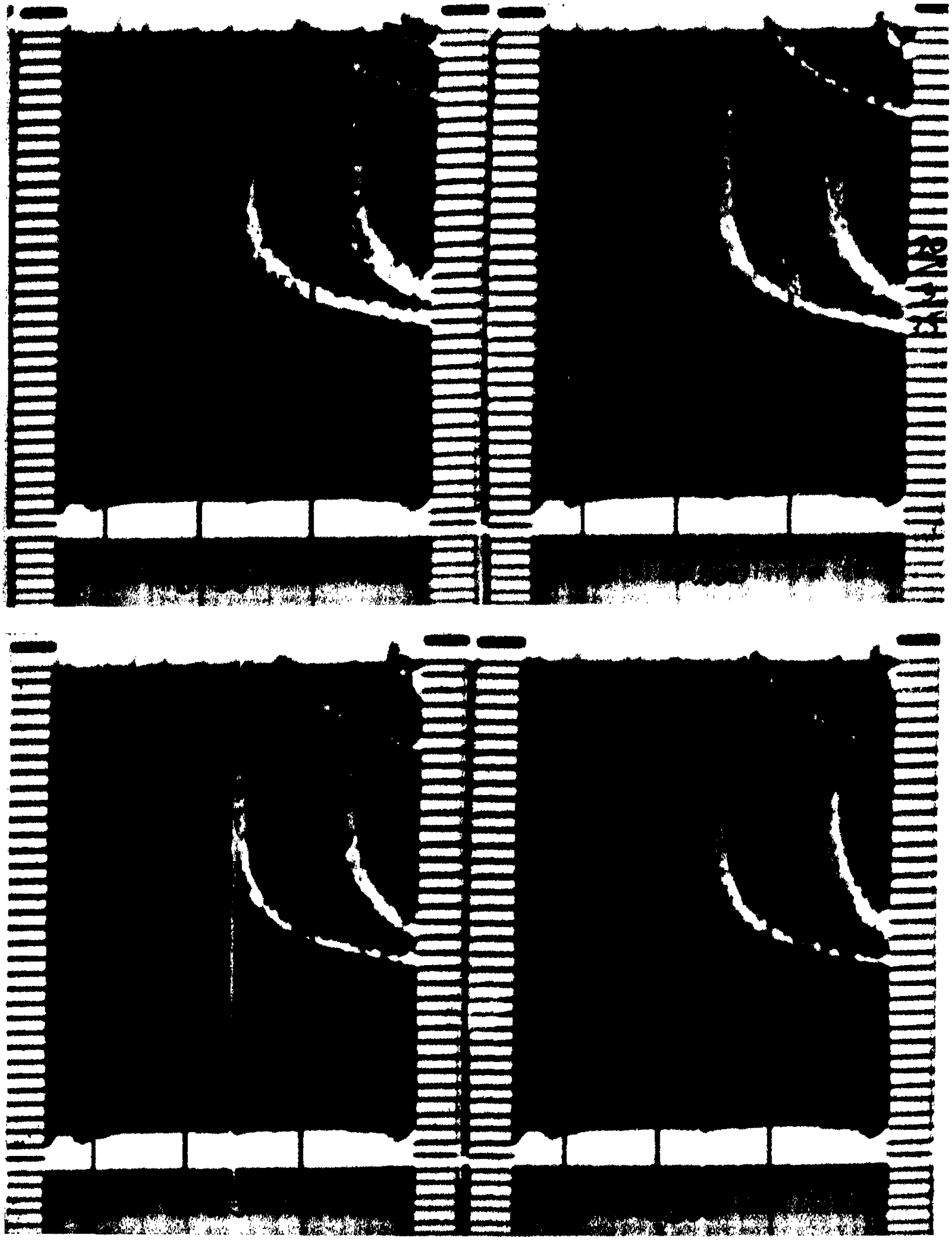


Fig. 6

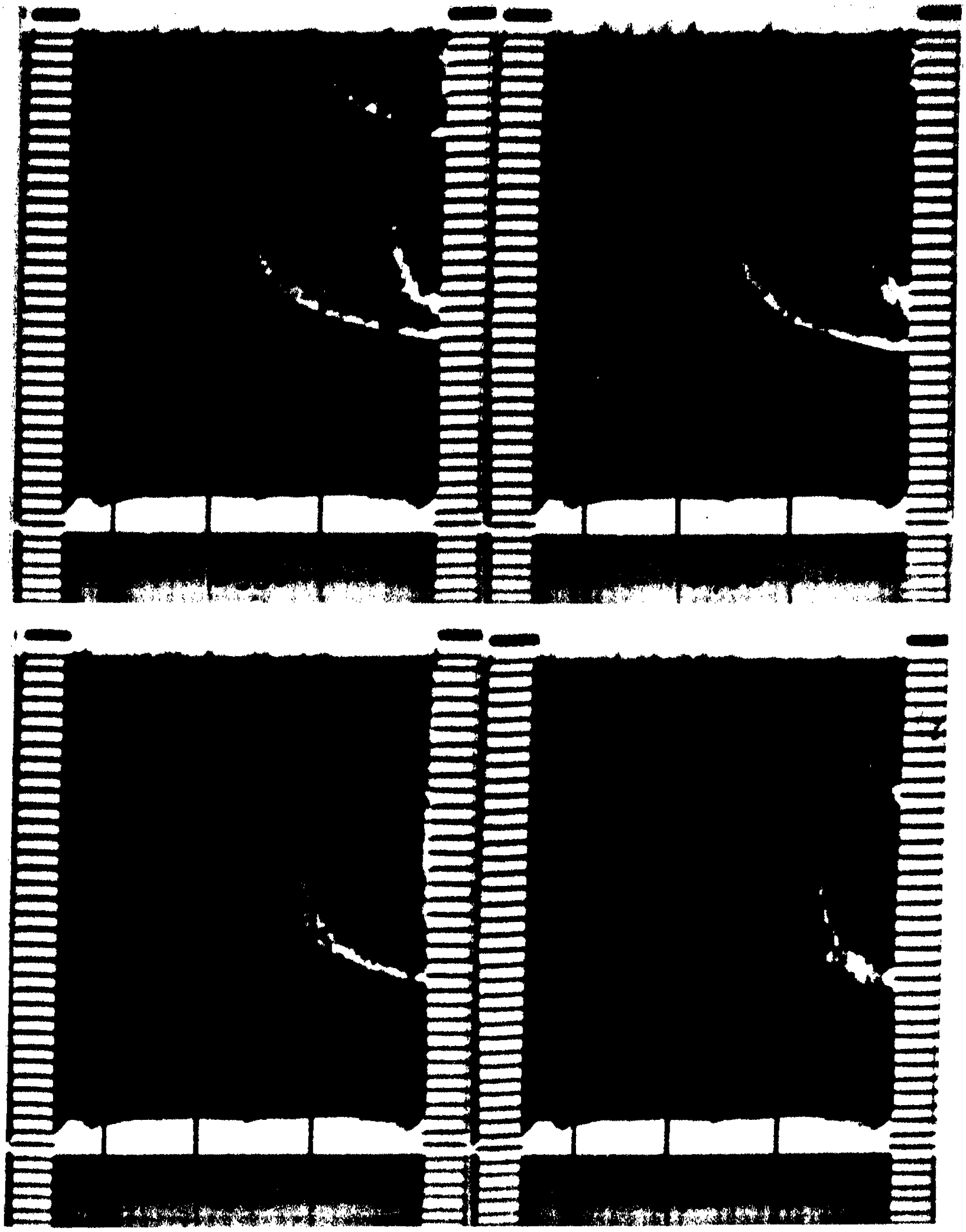


FIG. 6

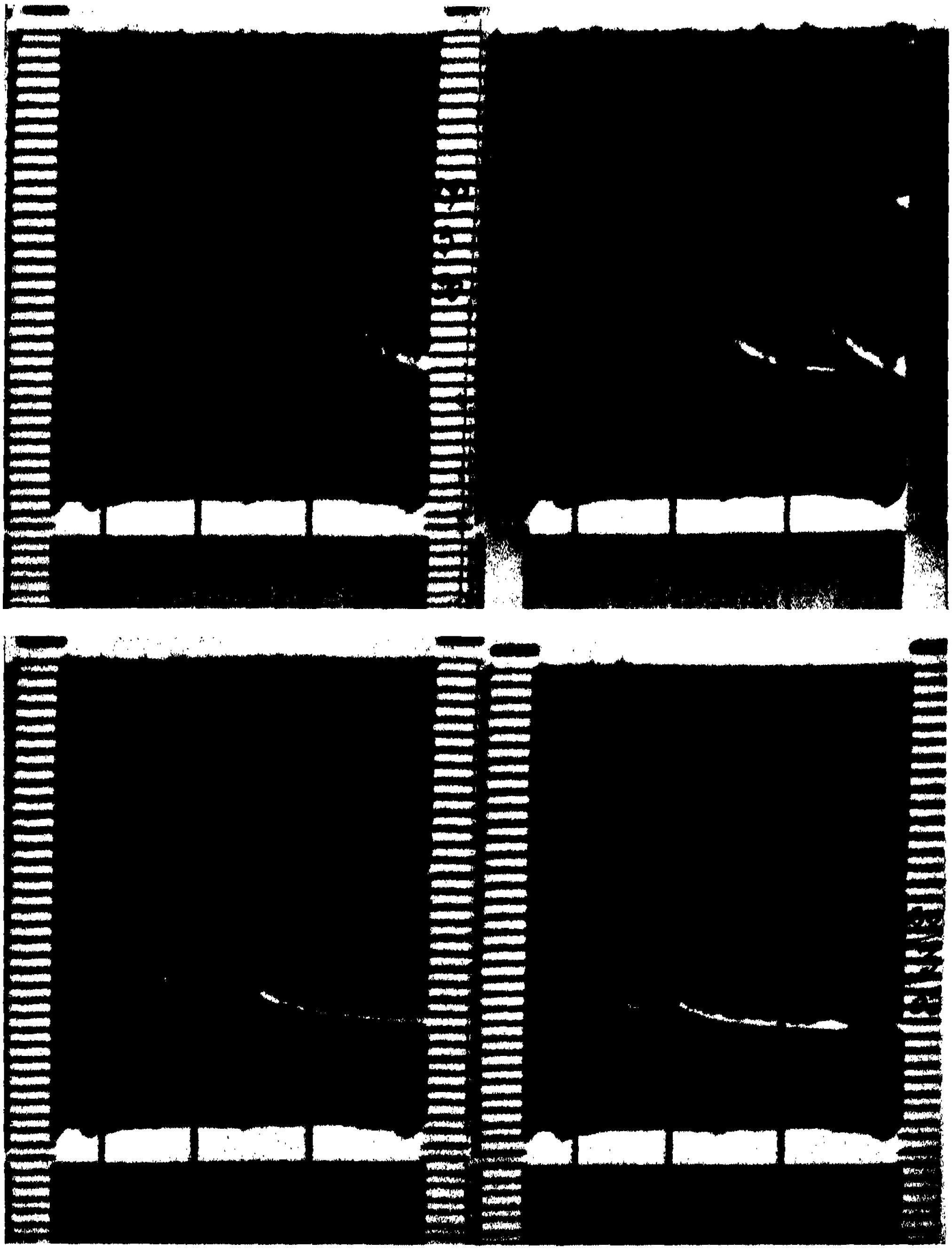


FIG. 6

conditions theory requires only partial dissociation of the oxygen molecules into atoms, whereas the spectroscopic evidence shows that such a dissociation is complete during the night and so must *a fortiori* be also complete during the day.

We therefore attempt to explain the observational results given above in terms of a recombination theory. In this connexion it should be noted that the rate of decay of electron concentration is given by $\alpha N_e N_+$ where α is the recombination coefficient, N_e is the electron concentration and N_+ is the positive ion concentration. We can thus write, for night-time conditions,

$$\frac{dN_e}{dt} = -\alpha N_e N_+, \tag{23}$$

or, if the medium is electrically neutral,

$$\frac{dN_e}{dt} = -\alpha N_e^2\left(1+\frac{N_-}{N_e}\right) = -\alpha' N_e^2, \tag{24}$$

when N_- is the negative ion concentration. The effective electron recombination coefficient thus depends on the ratio of the negative ion and electron concentrations. Now Massey has shown, in some unpublished work, that the processes of attachment and detachment of electrons to and from neutral atoms are such as to form an approximately constant balance between the negative ion and electron concentrations. By considering the magnitudes of the probability coefficients for these processes he concludes that for conditions in Region E this ratio is of the order of 100, which is in good agreement with the ratio suggested above from the comparison of the radio absorption and magnetic data.

It will be seen from (24) that the effective electron recombination coefficient α' is raised beyond the normal value α, which is due to the fact that N_+ is greater than N_e. With a ratio of 100 for N_-/N_e the value of α' can be of the order of 10^{-8} when α (see below) is of the order of 10^{-10}. The fact that α' is greater during the day than during the night is thus probably due to the greater value of N_-/N_e in the day-time.

Since an inferior limit has been estimated for the effective electron recombination coefficient in Region E, it is possible to give an inferior limit for the total rate of ion production in this region. Taking a summer day when the critical frequency for the ordinary ray is 3·5 Mc./sec. we find that the maximum electron concentration is $1{\cdot}5\times10^5$ electron per c.c. Since α is greater than 10^{-8} we see that q, the maximum rate of ion production, is greater than 225 c.c. per sec. Regarding $\cos\chi$ as approximately unity at summer noon, we find from (6) (taking H as 10^6 cm.) that an inferior limit for Q is $6 \cdot 1 \times 10^8$ electrons per sec./sq. cm. column for the region in question.

In the case of Region F_2 it is possible to measure the effective value of α both during the day and during the night. There is little doubt that the value found is less than that for the lower Regions E and F_1.

To illustrate this the values of α given in Table II may be quoted. They are derived from observations made during nights when a steady decay of N_e was noted, with no signs of nocturnal increase such as is noted in winter.

TABLE II

Period	α (Recombination coefficient)
25/26 April 1933	$3 \cdot 7$ to $6 \cdot 0 \times 10^{-10}$
11/12 April 1934	$3 \cdot 6 \times 10^{-10}$
April (average) 1934	$2 \cdot 5 \times 10^{-10}$
16/17 May 1934	$2 \cdot 9 \times 10^{-10}$
24/25 July 1934	$3 \cdot 4 \times 10^{-10}$
September (average) 1934	$1 \cdot 84 \times 10^{-10}$

The decay in nocturnal Region F ionization for the night 19/20 March 1935 is shown in fig. 6, Plates 23–28, in a series of equivalent height-frequency curves taken with the Slough automatic apparatus at half-hourly

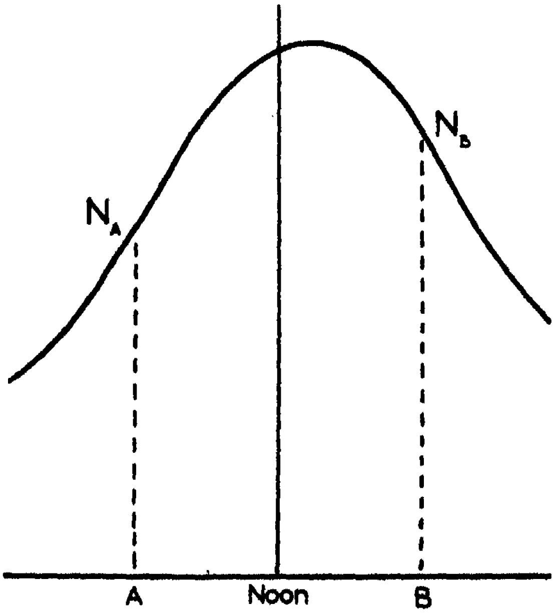


FIG. 7

intervals. The frequency scale runs from 2·6 Mc./sec. on the right to 5·2 Mc./sec. on the left, with frequency calibration marks at 3·26, 4·1 and 5·0 Mc./sec. respectively. Till 0400 both ordinary and extraordinary critical frequencies can be determined. From 0430 to 0530 only the extraordinary critical frequency is shown. Till 0530 the critical frequencies are gradually becoming lower. With the rise of ionization at 0600 the increase of critical frequencies begins. The analysis of these records shows that, on the night in question, α was $2{\cdot}8 \times 10^{-10}$.

There is, however, a method of determining both α and q during the day which is found to be applicable to Region F_2 where, owing to the relatively feeble recombination, the values of N are not symmetrically developed relatively to noon. Such an asymmetrical variation of N with time is shown in fig. 7. Consider in fig. 7 two times, A and B, equally spaced from noon. At these times the value of q, the electron production, is the same. Also, if there has been no atmospheric alteration, the level of maximum ionization will be at the same height. We have, therefore,

$$\frac{dN_A}{dt} = q - \alpha N_A^2 \tag{25}$$

and

$$\frac{dN_B}{dt} = q - \alpha N_B^2, \tag{26}$$

so that

$$\alpha(N_B^2 - N_A^2) = \frac{dN_A}{dt} - \frac{dN_B}{dt} \tag{27}$$

and

$$q = \frac{N_B^2 \dfrac{dN_A}{dt} - N_A^2 \dfrac{dN_B}{dt}}{N_B^2 - N_A^2}. \tag{28}$$

Using this method we may find q, and also α, at times A and B, and, knowing χ at time A or B, we can calculate the corresponding values at noon.

The application of this method to the case of a mid-summer day and an equinox day yields, for values of q at noon, 78 and 88 electrons per c.c. per sec. respectively and, for α, $8{\cdot}7 \times 10^{-11}$ and $8{\cdot}1 \times 10^{-11}$ c.c. per electron per sec. respectively. These values relate to south-east England. (The method is not found to be applicable to the case of a winter day when the ionization density variation is more symmetrical relative to noon, showing a higher effective electron recombination coefficient. This may be due to the fact that Region F_2 is then lower down, overlapping Region F_1, and so behaving like it as regards recombination.) It will therefore be seen that the value of α during the day is a little smaller than during the night.

The total region ion production Q in Region F_2 may be estimated with the aid of equation (6). Taking q as 80, and H as 50 km., we find Q to be $1{\cdot}1 \times 10^9$ electrons per sec. per sq. cm. column.

A comparison may be made between the experimentally obtained value of α (i.e. 3×10^{-10}) for night-time decay and the theoretical formulae for the recombination of positive ions and electrons. Chapman (1931c) has, for example, shown that Milne's formula for α gives a value of $3{\cdot}2 \times 10^{-10}$ when the absolute temperature is 400° and when Z, the atomic number, is taken at 8 for oxygen. Recombination to the ground state is presumably here contemplated. But we must take into account the possibility of recombination to excited states as well as to ground states. In the former case Z must be of the order of unity, while in the latter it is unlikely to be as large as 8; according to Slater's rules it is, for example, 3·7. This would reduce the value of α to 5×10^{-12} (excited state) and to $6{\cdot}9 \times 10^{-11}$ (ground state) for the temperature quoted. For higher temperatures the value will be slightly lower.

A theoretical value may be deduced in another way. Morse and Stueckelberg (1930) have shown that the cross-section $q(n, l)$ for recombination for slow electrons is given approximately by

$$q(n, l) = \frac{A(n, l)\, 10^{-20} Z^2}{V}, \tag{29}$$

where $A(n, l)$ depends on the orbit into which the electron is captured and V is the electron-volt energy.

Now the recombination coefficient is given by

$$\alpha = q(n, l)\sqrt{\frac{2eV}{m}}, \tag{30}$$

so that its value can be estimated, using the values of $A(n, l)$ given by Morse and Stueckelberg. Taking $T = 400°$ K. and thus $V = 0{\cdot}05$, we find for return to an excited state ($Z = 1$ and $A(n, l) = 0{\cdot}05$) that α is $1{\cdot}3 \times 10^{-13}$ and for return to the ground state ($Z = 3{\cdot}7$ and $A(n, l) = 0{\cdot}3$), $\alpha = 1{\cdot}1 \times 10^{-11}$. For higher temperatures corresponding values can be estimated, noting that α varies inversely as the square root of the absolute temperature. It thus appears that both theoretical formulae give a value for α which is smaller than the nocturnal value given above, which might suggest that negative ion formation played a part in raising the effective electron recombination coefficient in Region F as suggested for Region E. But the experiments of Kenty (1928) on the recombination of electrons and ions in argon, which does not form negative ions, gave a value of α which when corrected, to the temperature of 400° K. taken as an example above, is of the order of 8×10^{-10}, which is only slightly higher than the radio value. There thus appears to be a discrepancy between the experimental and theoretical values for α, the former being higher than the latter.

The fact that ions are not present in large quantities in Region F_2 shows that the gas or gases there possess little electron affinity and suggests a preponderance of nitrogen (if helium be not present). This would be in agreement with the fact that the auroral spectrum shows a greater intensity of the nitrogen band spectrum relatively to the auroral green line at the higher auroral levels. Such a state of affairs is curious, however, since if diffusive separation does take place at these high levels there ought to be a preponderance of the lighter oxygen atoms over the heavier nitrogen molecules.

(*e*) *The Actual Height reached by Radio Waves and the Structure of the Atmosphere*

The quantities usually measured in radio sounding are the equivalent heights of the reflecting strata for different frequencies. From these it is desired to deduce the relation between electron density and height. We now attempt to do this in two simple cases.

We consider a simple region formation of parabolic distribution and also neglect the influence of the earth's magnetic field, so that our deductions will apply approximately to the case of the ordinary ray propagation.

Let us take as the relation between N, the electron concentration, and y, the height (measured from the lower boundary of the region),

$$\frac{Ne^2}{\pi m} = f_c^2\left(2\frac{y}{y_m} - \frac{y^2}{y_m^2}\right), \tag{31}$$

where f_c is the critical frequency and y_m is the half-thickness of the stratum. We then have for the refractive index

$$\mu^2 = 1 - \frac{f_c^2}{f^2}\left(\frac{2y}{y_m} - \frac{y^2}{y_m^2}\right). \tag{32}$$

Under such conditions we can find the group path P', the optical path P, and the actual height reached by waves of any given frequency. These are as follows:

$$P' = y_m \frac{f}{f_c} \log\left(\frac{f_c+f}{f_c-f}\right), \tag{33}$$

$$P = y_m\left[1 - \frac{f_c^2-f^2}{2ff_c}\log\left(\frac{f_c+f}{f_c-f}\right)\right], \tag{34}$$

$$y = y_m\left[\frac{f_c - \sqrt{f_c^2-f^2}}{f_c}\right]. \tag{35}$$

These relations satisfy, as they should, the more general relations

$$P' = P + f\frac{\partial P}{\partial f} \tag{36}$$

and

$$y = \frac{1}{\pi}\int_0^f \frac{P'(\xi)}{(f^2-\xi^2)^{\frac{1}{2}}}\cdot d\xi. \tag{37}$$

In fig. 8 are plotted the equivalent height h' (i.e. $P'/2$), actual height of reflexion y, and optical height h (i.e. $P/2$) as functions of frequency. The actual values of the ordinates and abscissae chosen are such as to give a very close fit with the experimental relation between h' and f for 0000 h. on 20/6/34. We may thus conclude that at this particular hour of the night the ionization is fairly well represented by a region with a lower boundary at 270 km. and a half thickness (i.e. distance from under boundary to maximum) of 100 km.

If there were only one reflecting region in the upper atmosphere the case of $f \leqslant f_c$ would suffice, but in the case of the day-time reflexion by Region F we must take into account the influence of group retardation phenomena due to transmission through Region E. (This effect has been neglected in the nocturnal example taken immediately above.) Now when the frequency of the waves exceeds the critical frequency of a parabolic region, we have

$$P' = 2y_m\frac{f}{f_c}\log\left(\frac{f+f_c}{f-f_c}\right) \tag{38}$$

and

$$P = 2y_m\left[1 + \frac{f^2-f_c^2}{2ff_c}\log\frac{f+f_c}{f-f_c}\right], \tag{39}$$

while of course, the actual distance traversed in the double transmission is $4y_m$.

To illustrate the effect of group retardation we take the case of two regions with values of y_m and f_c chosen to fit conditions in mid-winter 1936. Good agreement between the values of equivalent height h' and frequency is shown when the values of y_m for Regions E and F are taken as 20 and 100 km. respectively.

The fact that different values of y_m for Regions E and F have to be used has considerable significance, for as fig. 1 shows, we should expect this quantity to be determined by the "scale height" H at the level in question. The results exhibited in fig. 9 leave little doubt that the value of H for the ionizable gases is some four or five times as large in Region F as in Region E. Comparing figs. 1 and 9 we see that H is of the order of 10 km. for Region

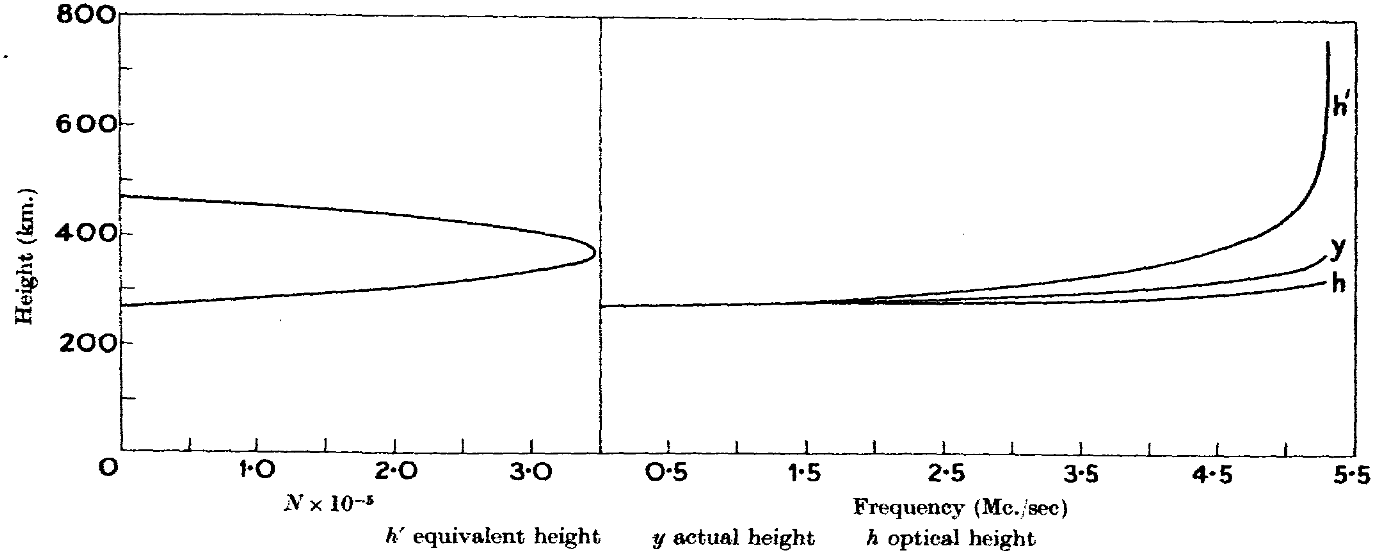


FIG. 8—Theoretical curves of equivalent, actual and optical heights as functions of frequency for a "parabolic" layer. The assumed ionization distribution is shown on the left.

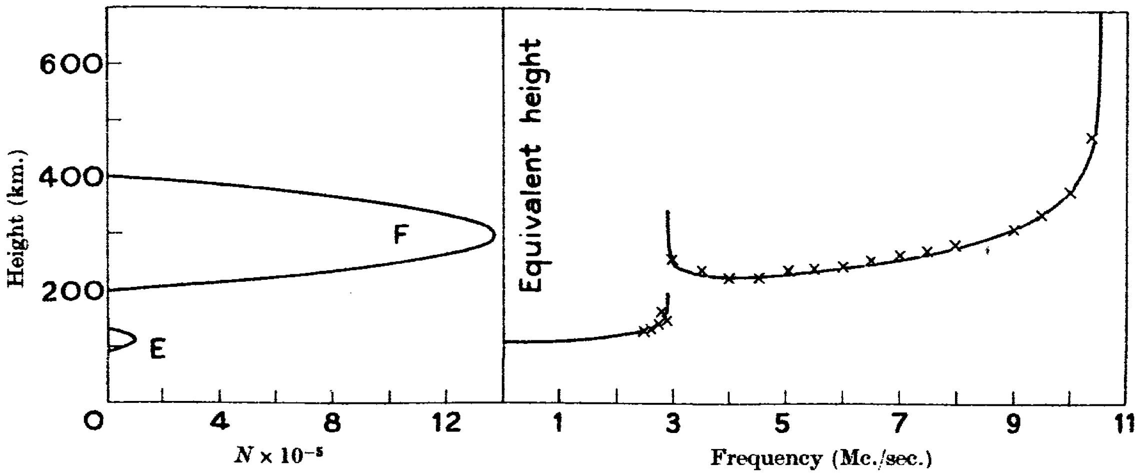


FIG. 9—Theoretical relation between equivalent height and frequency for two "parabolic" layers. The assumed ionization distribution is shown on the left.

E and 50 km. for Region *F*. We then see that some alteration in the structure of the atmosphere is indicated somewhere between Region *E* and Region *F* on a winter day. Martyn and Pulley have already drawn attention to the fact that the experimental values of electron collisional frequency do not (at levels beyond 100 km.) fall off as rapidly as might be expected if the molecular density from ground level is assumed to decay exponentially with height. The evidence presented above carries this matter farther in showing that the ionizable component (as distinct from the total molecular density) does not decay exponentially with height from ground level. There is, it appears, some influence at work which spreads the atmosphere out to high levels, and it is due to this influence that Region *F* can have ionization densities at levels of 300 km. as high as a million electrons per c.c.

Now from (2) we can see that an increase in the scale height *H* in passing from Region *E* to Region *F* can be brought about by an increased temperature, or by a decreased mean molecular mass, or by a combination of these. For the effect to be due simply to a decrease of mean molecular gas it is seen that the higher atmosphere must consist largely of helium, for a change to atomic oxygen at high levels would not give the increase in *H* required by the results.

In favour of the alternative explanation in terms of an increased temperature at high atmospheric levels it should be noted that spectroscopic indications of the presence of helium are yet lacking, while there are certainly not enough photons of the required energy in the sun's spectrum (radiating as a black body at 6000° K.) to produce the number of ion pairs required in a gas with such a high ionization potential as helium. Moreover, a high upper atmospheric temperature would explain the absence of helium since, like hydrogen, it would escape from the earth's gravitational field.

We may say, then, that the evidence available at present is more satisfactorily explained by the high-temperature hypothesis, but that the alternative explanation in terms of a light gas such as helium cannot be rejected with complete certainty.

3—ABNORMALITIES

The correlation between magnetic storms and long-wave long-distance transmission phenomena was first noted by the engineers of the Western Electric Company (Anderson 1928), who found that the accompaniment of a storm was a decreased night-time signal and an increased day-time signal. The corresponding correlation for short-wave transmission was given by Appleton (1927), who found that magnetic storms were associated with abnormally low values of signal intensity received from American short-wave stations in England. An attempt to reconcile the different effects noted for different wave-lengths was made by Appleton, who pointed out that, if a magnetic storm were accompanied by an increase in ionization in and below Region *E*, there would be an increase in the reflexion coefficient of this Region for long waves and yet a greater absorption of short waves which, penetrating Region *E*, were ultimately reflected by Region *F*.

The development of the critical frequency method of investigating ionospheric densities has made it possible to examine in greater detail the ionospheric abnormalities accompanying magnetic storms. The most marked connexion of this kind is the inverse correlation between magnetic activity and Region F_2 ionization. This was found by Appleton and Ingram (1935) in a comparison of ionospheric data and magnetic character figure supplied by the Abinger Observatory. In this comparison it was found that intense magnetic storms were usually accompanied by marked reductions in Region F_2 noon critical frequency, though there were noted a number of cases in which the critical frequency was actually greatly increased while a storm was in progress. The matter was investigated further by the evaluation of the correlation factor for the noon critical frequency of Region F_2 and the magnetic character figure of the preceding day. For the period August 1932–August 1933 this was found to be $-0{\cdot}4$. Later observations in this country as well as measurements made by Harang (1936) in Norway and by Kirby and others (1936) in America have amply confirmed this rather surprising inverse correlation.

The ionospheric abnormalities which accompany magnetic activity can best be illustrated by comparing the curves relating equivalent height of reflexion with frequency on quiet and disturbed days. Two such curves for the ordinary magneto-ionic component are shown in fig. 10. In this connexion it may be stated that, with disturbed magnetic conditions, the following radio phenomena are encountered:

(i) The critical frequency of Region F_2 is low and poorly defined.

(ii) Group retardation phenomena, more marked than normal, are associated with the penetration of Region F_1.

(iii) The Region F_1 critical frequency is not nearly so markedly altered as is that of Region F_2, but it is usually slightly reduced. (This reduction was first noted by Schafer and Goodall (1935).)

(iv) The minimum equivalent height of Region F_2 is very markedly raised.

(v) The radio echo amplitude is low for all frequencies employed (1·5 to 8 Mc./sec.).

Since 1935 a new type of ionospheric abnormality of an entirely different nature has been recognized. Its characteristics, so far as its practical effect on short-wave communication is concerned, were first noted by M. Jouaust, the chairman of the French National Committee of the International Scientific Radio Union. Through the offices of the Union M. Jouaust has organized a world-wide investigation of these phenomena from which the following points have emerged:

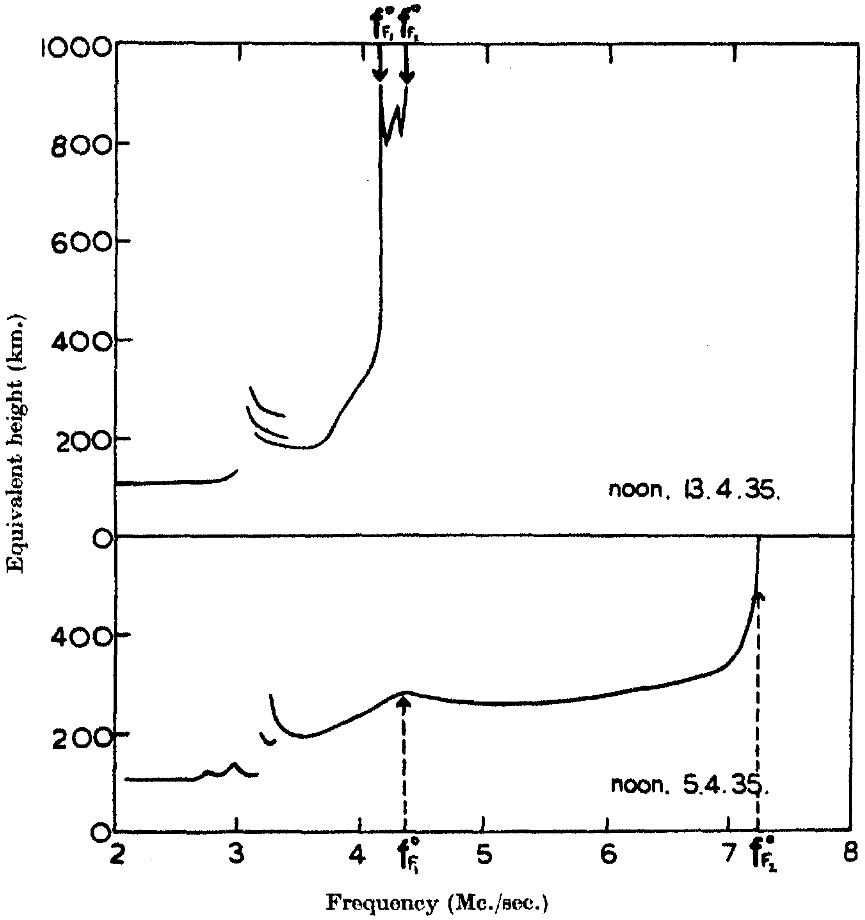


FIG. 10—Experimental equivalent height/frequency relation for magnetically disturbed day (13 April 1935) and for magnetically quiet day (5 April 1935).

(i) The abnormality is characterized by a sudden diminution in the intensity of short-wave long-distance signals. This may last a period of from 15 to 60 min.

(ii) Only circuits traversing the hemisphere illuminated by the sun are affected.

(iii) A number of these short-wave "fade-outs" have been found to occur at the time at which bright solar eruptions are observed by terrestial observers.

The radio stations at Slough and Cambridge have co-operated in the investigation of these sudden disturbances and it has been found that their characteristics when examined by vertical radio sounding are as follows. Before the disturbance occurs the ionosphere may be normal. During the disturbance there is a marked reduction in echo-intensity. After the disturbance the equivalent height/frequency relation is normal. There is thus a striking difference between this type of fade-out and the magnetic storm type described above. It appears, in fact, that the type of fade-out associated with bright eruptions produces extra ionization below the maximum of Region *E* (possibly in the absorbing layer dealt with in (2*a*; above) which causes attenuation of the radio waves. Such ionization, as we have seen, would be maintained at the equilibrium value and so would rapidly disappear when its ionizing source was removed.

ACKNOWLEDGEMENTS

The experiments described in this paper were carried out as part of the programme of the Radio Research Board of the Department of Scientific and Industrial Research. I wish to thank Miss L. J. Ingram, Mr. R. Naismith and Mr. W. R. Piggott for assistance in their prosecution. I also wish to express my indebtedness to Mr. R. D. Gander, my assistant at King's College, London, during the last twelve years, for his skilled and enthusiastic co-operation in ionospheric investigations conducted during that period.

REFERENCES

Anderson 1928 *Proc. Inst. Radio Engrs, N.Y.*, **16**, 297.
Appleton 1927 *Electrician* (11 March), p. 256.
Appleton and Ingram 1935 *Nature, Lond.*, **136**, 548.
Appleton and Naismith 1935 *Proc. Roy. Soc.* A, **150**, 685.
Burrows 1931 *Proc. Inst. Radio Engrs, N.Y.*, **19**, 1634.
Chapman, S. 1931*a* *Proc. Phys. Soc.* **43**, 26.
— 1931*b* *Proc. Phys. Soc.* **43**, 484.
— 1931*c* *Proc. Roy. Soc.* A, **132**, 352.

Eckersley 1935 *Nature, Lond.*, **135**, 435.
Farmer and Ratcliffe 1935 *Proc. Roy. Soc.* A, 151, 370.
Harang 1936 *Beitr. Geophys.* **46**, 438.
Kenty 1928 *Phys. Rev.* **32**, 624.
Kirby, Gilleland, Smith and Rayner 1936 *Phys. Rev.* **50**, 258.
Lenard 1911 *S.B. Akad. Wiss. Heidelberg*, **12**.
Martyn and Pulley 1936 *Proc. Roy. Soc.* A, **154**, 455.
Morse and Stueckelberg 1930 *Phys. Rev.* **36**, 16.
Pekeris 1937 *Proc. Roy. Soc.* A, **158**, 650.
Saha 1937 *Proc. Roy. Soc.* A, **160**, 155.
Schafer and Goodall 1935 *Proc. Inst. Radio Engrs, N.Y.*, **23**, 670.
White and Brown 1936 *Proc. Roy. Soc.* A, **153**, 639.

DESCRIPTION OF PLATES 23-28

FIG. 6—(P', f) records taken at half-hourly intervals at Slough Radio Research Station during the night 19/20 March 1935. The ordinate scale is marked off at 25 km. intervals.

The Tests for Sampling Differences and Contingency

BY HAROLD JEFFREYS, F.R.S.

(*Received* 22 *June* 1937)

1—The question discussed in these tests (Jeffreys 1935, 1936*a*) has the form: we have a sample or a pair of samples, in which the numbers of individuals with the pairs of properties $\phi . \psi$, $\phi . \sim\psi$, $\sim\phi . \psi$, $\sim\phi . \sim\psi$ are respectively x, y, x', y': do the numbers afford evidence for or against the proposition that ϕ and ψ are associated? The conditions contemplated in the experiments, however, are somewhat different. In what I have called the sampling problem (1935, p. 203), large classes of ϕ's and $\sim\phi$'s are already supposed separated, and we select from them at random arbitrarily assigned numbers of members $x+y$, $x'+y'$. In the contingency problem (1936*a*, pp. 426–29), the ϕ's and $\sim\phi$'s are not already separated. We have only a single large class containing all four types, and in the sampling we choose arbitrarily a total number $x+y+x'+y'$. It would therefore be expected that the significance tests found in the two cases would differ slightly. In the sampling problem, the sampling errors of x and y are constrained to be equal and opposite, which is not so in the contingency problem. The tests actually found were different, and no surprise was felt; it is only recently that I have noticed that the difference is liable to be so large that in some cases, with the same x, y, x', y', an association could not be asserted in the contingency problem when it would be asserted without hesitation in the sampling one. Again, if x' and y are very different, interchanging them makes a great difference to the result, so that in the sampling problem it apparently makes a large difference whether we sample first with regard to ϕ or ψ if one of them is a rare property. It turns out that the discrepancy is removed when we take into account an inequality that was overlooked.

We denote the proposition that there is no association by q, that there is one by $\sim q$. K is the ratio expressing the support for q given by the data θ, namely

$$K = \frac{P(q \mid \theta h)}{P(\sim q \mid \theta h)} \Big/ \frac{P(q \mid h)}{P(\sim q \mid h)}. \tag{1}$$

Case I is the case of sampling where $x+y$ ϕ's and $x'+y'$ $\sim\phi$'s are selected and then counted with regard to ψ; Case II is the analogous case except that $x+x'$ ϕ's and $y+y'$ $\sim\phi$'s are chosen and then counted with regard to ψ; Case III is the contingency case, where a sample of number $x+x'+y+y'$ is chosen and then counted with regard to both ϕ and ψ. The corresponding values of K, according to the previous work, would be:

$$\text{Case I:} \quad K_1 = \frac{(x+x')!\,(y+y')!\,(x+y+1)!\,(x'+y'+1)!}{x!\,y!\,x'!\,y'!(x+y+x'+y'+1)!}; \tag{2}$$

$$\text{Case II:} \quad K_2 = \frac{(x+y)!\,(x'+y')!\,(x+x'+1)!\,(y+y'+1)!}{x!\,y!\,x'!\,y'!\,(x+y+x'+y'+1)!}; \tag{3}$$

$$\text{Case III:} \quad K_3 = \frac{(x+y)!\,(x'+y')!\,(x+x')!\,(y+y')!\,(x+y+x'+y'+3)!}{3!\,x!\,y!\,x'!\,y'!\,\{(x+y+x'+y'+1)!\}^2}. \tag{4}$$

If ϕ is a rare property, so that with x and y moderate x' and y' are large in Cases II and III, we write N for the sum of all four and consider in Case I a pair of samples such that $x+y$, $x'+y'$ are the same as in the other cases. Then if the separate numbers are the same we have

$$\frac{K_2}{K_1} = \frac{(x+x'+1)\,(y+y'+1)}{(x+y+1)\,(x'+y'+1)} = O\left(\frac{N}{x+y}\right), \tag{5}$$

$$\frac{K_3}{K_1} = \frac{(N+2)\,(N+3)}{6(x+y+1)\,(x'+y'+1)} = O\left(\frac{N}{x+y}\right), \tag{6}$$

and these ratios are large numbers, of the same order of magnitude. Thus with the same numbers in the three cases K_1 may be much less than 1, supporting an association, when K_2 and K_3 are much larger than 1 and deny any association. It is plainly unsatisfactory that a mere redistribution of the randomness due to the conditions of sampling should make such a difference.

1·1—The easiest case to consider further is Case III, on account of its symmetry and the reduced arbitrariness of the sizes of the samples. In this case, we take the chances of a specimen having the four possible pairs of properties as $p_{11}, p_{12}, p_{21}, p_{22}$; then these were taken previously to be mutually irrelevant except for their sum being 1. In these conditions the prior probability distribution is

$$P(dp_{11}dp_{12}dp_{21} \mid \sim q.h) = 3!\, dp_{11}dp_{12}dp_{21}. \tag{7}$$

Now the assumption of mutual irrelevance is not obvious even if there is contingency. If p_{11} is small it means that $\phi\psi$ is a rare combination of properties. This is capable of three interpretations: ϕ or ψ may be rare, or both may be common but rare in conjunction. With any moderate degree of contingency, the smallness of p_{11} in the first two cases implies that of either p_{12} or p_{21}, and therefore that p_{22} is large. The third case would express so high a degree of contingency that it would be revealed by almost any test.* We are really concerned with the first two. But if small p_{11} establishes a moderate probability that p_{12} is also small the hypothesis of prior irrelevance is wrong.

A better procedure appears to be to regard the problem of contingency as an attempt to answer three questions. How common is the property ϕ? How common is ψ? Are ϕ and ψ associated? The third question then does not arise until we have answers to the first two, since the values of p_{11} and so on, on the hypothesis of no association, cannot be estimated until we know the chances of ϕ and ψ separately. These chances, however, have a meaning whether there is contingency or not. Thus if a is the chance of ϕ, a' of $\sim\phi$, b that of ψ, b' of $\sim\psi$, then on the hypothesis q of no association we shall have $p_{11} = ab$ and so on, with $a+a' = 1$, $b+b' = 1$. On hypothesis

---

* The case of $p_{11}=0$ may in fact have a finite prior probability, as for the problem of simple sampling, in which the commonness of cases where all members of the class sampled have the property under discussion leads to the result that Laplace's uniform assessment of the prior probability no longer holds for the extreme cases, if it ever did (Jeffreys 1937*a*, p. 253). The appropriate modification would not be difficult, but we are dealing here with cases where specimens of all four types have actually occurred, so that $p_{11}=0$ is excluded by the data, and needs no special attention.

$\sim q$ there will be a further parameter c such that the chances of the various conjunctions are

$$\begin{pmatrix} \phi\psi & \phi.\sim\psi \\ \sim\phi.\psi & \sim\phi.\sim\psi \end{pmatrix} \vdots\vdots \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} ab+c & ab'-c \\ a'b-c & a'b'+c \end{pmatrix}, \tag{8}$$

where a and b retain the same values. Further, a and b are cases where Laplace's theory applies, since the extreme cases $a, b = 0$ or 1 are excluded. Hence

$$P(dadb \mid qh) = P(dadb \mid \sim qh) = dadb. \tag{9}$$

Proposition q asserts that $c = 0$. On $\sim q$ we find easily

$$p_{11}p_{22} - p_{12}p_{21} = c, \tag{10}$$

$$\frac{\partial(p_{11}+p_{12}, p_{11}+p_{21}, p_{11}p_{22}-p_{12}p_{21})}{\partial(p_{11}, p_{12}, p_{21})} = 1. \tag{11}$$

Hence the uniform distribution of the probabilities of p_{11}, p_{12}, p_{21} within the permissible ranges is compatible with that for a, b, c. There is no displacement of the maximum posterior density. Here, however, the permissible values of c are conditional on those of a and b, since all the p_{st} must lie between 0 and 1. We can suppose the diagram so arranged that $a \leqslant a'$, $b \leqslant b'$; then from the first line in (8) c must be between $-ab$ and $+ab'$. Within this range p_{22} is always possible; and p_{21} is always possible provided $a'b \geqslant ab'$. The condition that this can be the range of possible values of c is then that $ab, ab', a'b, a'b'$ shall be an increasing sequence. This can always be achieved by a permissible rearrangement of the diagram. The range of possible values of c is then a, and we have, within the permitted range,

$$P(dc \mid a, b, \sim q, h) = dc/a, \tag{12}$$

$$P(da\,db\,dc \mid \sim q, h) = da\,db\,dc/a. \tag{13}$$

Also, as in the previous analysis for contingency,

$$P(\theta \mid a, b, q, h) = a^{x+y}\, a'^{\,x'+y'}\, b^{x+x'}\, b'^{\,y+y'}, \tag{14}$$

$$P(\theta \mid a, b, c, \sim q, h) = p_{11}^{x} p_{12}^{y} p_{21}^{x'} p_{22}^{y'}. \tag{15}$$

These are the probabilities of getting the observations in their actual order. If only the total numbers of each type are known, they must be multiplied by the number of ways of dividing $x+y+x'+y'$ things into four groups of numbers x, y, x', y'; but this factor cancels from the posterior probability. Then

$$P(q \mid \theta h) \propto \int_0^1 \int_0^1 a^{x+y} a'^{x'+y'} b^{x+x'} b'^{y+y'} \, da \, db$$

$$\propto \frac{(x+y)!\,(x'+y')!\,(x+x')!\,(y+y')!}{(N+1)!\,(N+1)!}, \tag{16}$$

$$P(\sim q \mid \theta h) \propto \int_0^1 \int_0^1 \int_{-ab}^{ab'} \frac{1}{a} p_{11}^{x} p_{12}^{y} p_{21}^{x'} p_{22}^{y'} \, da \, db \, dc. \tag{17}$$

By (11) we can replace $da\,db\,dc$ by $dp_{11}\,dp_{12}\,dp_{21}$ within the range permitted. The integrand has a sharp maximum, and without sensible error, so long as the maximum for c does not fall near the extreme values (when association will be obvious), the integral will be the same as over all values such that $p_{11}+p_{12}+p_{21} \leqslant 1$. Also near the maximum a is nearly $(x+y)/N$. Hence, approximately,

$$P(\sim q \mid \theta h) \propto \frac{N}{x+y} \frac{x!\,y!\,x'!\,y'!}{(N+3)!}, \tag{18}$$

$$\frac{P(q \mid \theta h)}{P(\sim q \mid \theta h)} = \frac{(x+y)!\,(x'+y')!\,(x+x')!\,(y+y')!}{x!\,y!\,x'!\,y'!\,N!}(x+y)$$

$$= \left\{\frac{N^3(x+y)}{2\pi(x+x')(y+y')(x'+y')}\right\}^{\frac{1}{2}} \exp\left[-\tfrac{1}{2}\frac{N(xy'-x'y)^2}{(x+x')(y+y')(x+y)(x'+y')}\right] \tag{19}$$

nearly.

The effect of this revised analysis is that a factor $N/6(x+y)$ has been removed. Now one way of stating the principle of inverse probability is that in comparing two hypotheses we choose the one that requires the less remarkable coincidence to give the observed data. In the case that led to this investigation x and y are both small in comparison with $x'+y'$. To explain this requires only the one coincidence, that ϕ should be a rare property, whether it is associated with ψ or not. But on the previous analysis p_{11} and p_{12} were taken as independent on the hypothesis of association, so that the smallness of x and y required two coincidences, which is not the case. The change is due to the allowance for this fact.

1·2—We return now to Case I. $x+y$ and $x'+y'$ are now assigned in advance. If there is no difference between the classes sampled the probabilities of ψ given ϕ and $\sim\phi$ are the same; hence we can introduce a quantity b, the same for both classes, to express this probability. Also

$$P(db \mid qh) = db, \tag{20}$$

$$P(\theta \mid bqh) = b^{x+x'}(1-b)^{y+y'}. \tag{21}$$

If there is a difference the respective probabilities of a ψ are p and p'; but we can still define b so that Nb will be the expectation of ψ's in a pair of samples consisting of $x+y$ ϕ's and $x'+y'\sim\phi$'s. Then

$$Nb = (x+y)p+(x'+y')p'. \tag{22}$$

Both p and p' must lie between 0 and 1. Within the permissible range for p, p' for a given b can have values from $Nb/(x'+y')$ to $(Nb-x-y)/(x'+y')$. But the most probable value of Nb will be nearly $x+x'$; the former value will then be permissible if $x<y'$ and the latter if $x'>y$, and if these inequalities are satisfied there is no further restriction. Then

$$P(db \mid \sim q,\, h) = db, \tag{23}$$

$$P(dp \mid b,\, \sim q,\, h) = dp, \tag{24}$$

$$\frac{\partial(b,p)}{\partial(p',p)} = \frac{x'+y'}{N}, \tag{25}$$

$$P(\theta \mid p, p', \sim q,\ h) = p^x(1-p)^y p'^{x'}(1-p')^{y'}, \tag{26}$$

$$P(q \mid \theta h) \propto \int_0^1 b^{x+x'}(1-b)^{y+y'}\,db = \frac{(x+x')!\,(y+y')!}{(N+1)!}, \tag{27}$$

$$\begin{aligned} P(q \mid \sim\theta h) &\propto \int_0^1\int_0^1 p^x(1-p)^y p'^{x'}(1-p')^{y'}\,db\,dp \\ &= \frac{x'+y'}{N}\iint p^x(1-p)^y p'^{x'}(1-p')^{y'}\,dp\,dp' \\ &\doteqdot \frac{x'+y'}{N}\,\frac{x!\,y!}{(x+y+1)!}\,\frac{x'!\,y'!}{(x'+y'+1)!}, \end{aligned} \tag{28}$$

$$\frac{P(q \mid \theta h)}{P(\sim q \mid \theta h)} \doteqdot (x+y+1)\frac{(x+y)!\,(x'+y')!\,(x+x')!\,(y+y')!}{x!\,y!\,x'!\,y'!\,N!}, \tag{29}$$

which is indistinguishable from (19) to the accuracy of the approximations, subject to x, y, x', y' being arranged as an increasing sequence in the diagrams.

1·3—If $x' < y$ (more strictly, if $(x'+y')p' < (x+y)(1-p)$) the possible values of p' impose a further restriction. p can no longer reach unity, its greatest possible value being $Nb/(x+y)$. Then (22) and (23) still hold, but

$$P(dp \mid b, \sim q, h) = \frac{x+y}{Nb}dp. \tag{30}$$

The limits in (28) are slightly altered, but in the important case

$$P(\sim q \mid \theta h) \propto \frac{x+y}{N}\frac{x'+y'}{N}\iint \frac{1}{b}p^{x}(1-p)^{y}p'^{x'}(1-p')^{y'}\,dp\,dp', \tag{31}$$

and at the maximum of the integrand $b = (x+x')/N$ nearly. Hence

$$\frac{P(q \mid \theta h)}{P(\sim q \mid \theta h)} \propto (x+x')\frac{(x+y)!\,(x'+y')!\,(x+x')!\,(y+y')!}{x!\,y!\,x'!\,y'!\,N!} \tag{32}$$

nearly; and this is obtained from (19) and (29) by simply interchanging x' and y according to the altered sign of their difference. Thus the appropriate modification of Case II gives a result equivalent to that of Case I. It is unimportant whether we sample first with regard to ϕ or ψ, or to both at once, provided that of the four sums $x+y$, $x+x'$, $y+y'$, $x'+y'$ the smallest is the one to go into the numerator of the outside factor.

To obtain compact answers it has been necessary to proceed to approximation at an early stage; this may be the explanation of the slight difference between $x+y$ in (19) and $x+y+1$ in (29), and other minor differences have been neglected at intermediate stages. Minor differences between the results were to be expected, as already explained, on account of the different relations imposed between the sampling errors, but it is satisfactory that the results now agree within the amounts that can be reasonably attributed to these differences.

As a numerical illustration we may take the data quoted by Fisher (1936, p. 99) from Lange on the convictions of twin brothers or sisters (of like sex) of convicted criminals. The table, arranged to satisfy the necessary inequalities, is as follows:

	Monozygotic	Dizygotic
Convicted	10	2
Not convicted	3	15

Then the factorial expression in (19) is

$$\frac{12!\;18!\;13!\;17!}{10!\;2!\;3!\;15!\;30!}\cdot 12 = \tfrac{1}{185}.$$

The less accurate exponential approximation gives $\frac{1}{189}$, showing that it may be used even at rather extreme degrees of association such as this. What the result shows is that, starting without information about whether there is any difference in criminality between monozygotic and dizygotic twins of criminals, we can assert on the data that the odds on the existence of a difference are 185 to 1. Fisher infers a difference on the ground that $P(\chi^2)$ on the hypothesis of no difference is $1/76500$.

1·4—In all these cases the complexity of the integrals prevents an exact answer from being obtained in a compact form. Such a solution can, however, be obtained when a sampling ratio is tested against a theoretical value. This has been regarded as the extreme form of Case I, when x' and y' tend to infinity in a finite ratio; p' is here already specified. As the solution for Case I has turned out to be only an approximation a direct treatment of this case becomes desirable. On hypothesis q, the probability of a ψ given ϕ is the known fraction p'; on $\sim q$ it is p, which may have any value from 0 to 1. The observed numbers of ψ's in a sample of $x+y$ ϕ's are x, y. Then

$$P(q \mid h) = P(\sim q \mid h) = \tfrac{1}{2}; \quad P(\sim q . dp \mid h) = \tfrac{1}{2} dp, \tag{33}$$

$$P(x, y \mid qh) = p'^x (1-p')^y, \tag{34}$$

$$P(x, y \mid \sim q, p, h) = p^x (1-p)^y, \tag{35}$$

$$P(q \mid x, y, h) \propto p'^x (1-p')^y, \tag{36}$$

$$P(\sim q \mid x, y, h) \propto \int_0^1 p^x (1-p)^y \, dp \propto \frac{x!\,y!}{(x+y+1)!}, \tag{37}$$

and finally
$$\frac{P(q \mid x, y, h)}{P(\sim q \mid x, y, h)} = \frac{(x+y+1)!}{x!\,y!} p'^x (1-p')^y. \tag{38}$$

When x and y are large an immediate application of Bernoulli's approximation gives

$$\frac{P(q \mid x, y, h)}{P(\sim q \mid x, y, h)} \doteqdot \left\{\frac{x+y}{2\pi p'(1-p')}\right\}^{\frac{1}{2}} \exp\left\{-\frac{\{x-p'(x+y)\}^2}{2(x+y)p'(1-p')}\right\}. \tag{39}$$

The latter is easily seen to be the limit of the more general results for x' and y' large, and both (38) and (39) are the limits of the results given for Case I previously when x' and y' tend to infinity in the ratio $p'/(1-p')$.

The following table indicates how the support varies with x and y for the simple case of testing an even chance when the sample is small. In this case $p' = \frac{1}{2}$, and the ratios are denoted by K.

x	y	K	x	y	K	x	y	K
1	0	1	1	1	$\frac{3}{2}$	2	2	$\frac{15}{8}$
2	0	$\frac{3}{4}$	2	1	$\frac{3}{2}$	3	3	$\frac{35}{16}$
3	0	$\frac{1}{2}$	3	1	$\frac{5}{4}$	4	4	$\frac{315}{128}$
4	0	$\frac{5}{16}$	4	1	$\frac{15}{16}$	5	5	$\frac{693}{256}$
5	0	$\frac{3}{16}$	5	1	$\frac{21}{32}$			

None of these results is particularly decisive, and most of them have little *weight in* Keynes's *sense* (1921, Chap. 6) *that the probabilities, with a little more information, might differ appreciably*. It is not till $x = 7, y = 0$ that K falls below 0·1, and not till $x = y = 80$ that it reaches 10. To get this amount of support for an even chance therefore needs about as much information as would fix a ratio found by sampling with a standard error of 4 %. The result unity for $x = 1, y = 0$ of course expresses the fact that a single specimen must be of one type or the other and therefore tells us nothing about the existence of bias. For an analogous reason the result for a given x is exactly the same for $y = x$ and $y = x + 1$.

If a sample should have the composition (7, 0) it would be inferred that p was nearly unity, and the possibility that it is exactly unity would begin to be considered. This would be the type of case that can arise in a Mendelian experiment where an even chance might be expected but actually all individuals of one type die in the egg. Such an unexpected general rule is of course not covered by my assessments of prior probabilities. I have simply suggested that as a practical working rule *all* hypotheses *stated* in a sufficiently definite form to have their consequences worked out should have the same prior probability. It would, however, be absurd to apply it to such a case as this, where the rule would not begin to be considered seriously until numerous verifications of it had already been discovered. A workable, though rough, rule would be to take the probability as $\frac{1}{2}$ at the actual stage when it begins to be taken seriously; another would be to assess a prior probability on Laplace's theory from the results of previous experiments of the same type, as already suggested by Pearson. Such cases are instances of the rise into importance of the "excluded middle", the status of which in inductive logic has been discussed elsewhere (Jeffreys 1936*b*, p. 346).

2—*Significance Tests in General*—The tests examined and revised above are particular cases of those for multiple sampling and contingency already given (1936*a*, pp. 428–32). The latter therefore also require revision. This is indeed almost obvious on inspection. The outside factors contain factors $N^{\frac{1}{2}(m-1)(n-1)}$, where N is the whole number of specimens and the table has m rows and n columns. The exponential factor is $\exp(-\frac{1}{2}\chi^2)$ and for random occurrences χ^2 will be $k \pm \sqrt{(2k)}$, where $k = mn - m - n + 1$. Thus the result is the product of an enormous factor and a small one, the latter itself being capable of large random variations when there are many degrees of freedom. It will therefore in general be either very large or very small, and the random variations may determine which. There is no provision for indecisive results, such as we may expect to occur often in practice. The explanation is that the method used compared only the two extreme cases of complete proportionality and complete independence among the chances; in addition the statement of the latter alternative is wrong for the reasons given above. In normal cases there are several degrees of freedom, of which some will be relevant to the chances and others not, and the method does not distinguish between these. It is to be expected that the result of comparing the extreme cases will give a decisive answer one way or the other, but in most cases some intermediate case is more probable on the data than either, and we cannot identify the relevant degrees of freedom without a closer analysis that will test one of them at a time. This can, I think, usually be done by an extension of the theory for simple sampling or contingency. If abnormality anywhere is as likely as not, then out of k members the prior probability that any particular one is abnormal is nearly $\log_e 2/k$ (1936*a*, p. 430). Thus we can test the extreme departure from proportionality by using the simple theory, but taking $P(q \mid h)/P(\sim q \mid h) = k/\log_e 2$ instead of 1. If this test adjudges its sampling ratio to be the result of random variation we can accept the whole as random. If it indicates an abnormality we can proceed to test other members, the prior probability now being given by Pearson's extension of Laplace's theory. This method has already been used twice (1936*a*, p. 441; 1937*b*, p. 39). A more accurate one may be possible, but would be much more complicated and would usually lead to the same results.

The actual value of χ^2 for the whole set of data now ceases to have much direct importance, but its evaluation is nevertheless desirable in practice. The entry that makes the largest contribution to χ^2 would naturally be the first to test by the above method. If χ^2 falls within the limits indicated by its standard error, as it does in about two cases out of three when the variation is in fact random, the variation may be at once accepted as random if there is no obviously outstanding contribution and the signs of the residuals show no systematic variation. The latter point may be tested as follows. When previous considerations arrange the groups in some definite order (e.g. in studying the variation of mortality with age) a persistence of sign in consecutive residuals is as likely as a change if the variation from the trial values is random. The numbers of persistences and changes can therefore be compared with the hypothesis that the chance of a persistence is $\frac{1}{2}$.* Similarly a residual of three times its standard error, if it is not specially selected, is significant if it is based on less than about 10,000 observations and can usually be accepted at sight. I have found these simple tests adequate in the great majority of cases in recent work on the times of transmission of earthquake waves to different distances. It is a drawback of the use of χ^2 by itself that it attends only to the magnitude of the residuals and not to their signs. This extra test, however, often indicates a systematic variation when χ^2 by itself is sufficiently greater than $n+\sqrt{(2n)}$ to suggest one but not to establish it. It is only in intermediate cases that any test more detailed than these three together is needed.

When a set of data is arranged in numerous groups there is a further difficulty about the use of χ^2 in its usual form. If there are, say, 32 degrees of freedom, the expected χ^2 on the hypothesis of randomness is 32 ± 8. If then one group contributes 16 to χ^2 and the rest only 24, the total is 40, which as it stands would be consistent with random variation. But the exceptional residual is four times its standard error, and if tested directly, even with allowance for selection, would practically always be taken as systematic. With a large number of groups the random error of χ^2 may easily hide a systematic variation that would be detected if it was tested directly.

Similarly, when previous considerations put the groups in some definite order, several consecutive groups can be combined and the totals compared. This reduces the random error of χ^2 considerably, while its effect on the systematic difference, if there is one, is small; thus the sensitiveness of the test is much improved. The same applies if the groups naturally fall into two classes. In fact it appears that our best way of detecting systematic differences from a suggested rule is always to arrange the work so as to ask and answer one question at a time.† Thus William of Ockham's rule "Entities are not to be multiplied without necessity", for quantitative use must be amplified to "Variation is to be taken as random until the contrary is shown; and new parameters expressing systematic differences, when they are suggested, must be tested one at a time unless there is specific reason to the contrary". As examples of a specific reason we may mention the cases of a suggested periodic variation of unknown phase, where coefficients of a cosine and sine enter the solution simultaneously, and of two earthquake epicentres tested for identity, where if there is a difference in latitude there would naturally also be one in longitude.*

---

* If a straight line is fitted to the data by least squares, the residuals must necessarily have two changes of sign. The changes must then be reduced by 2 before applying the test.

† Fisher (1936, pp. 68, 106, 291) recommends a partition of χ^2 among the degrees of freedom, testing each separately. This is a considerable step in the present direction.

This does not say that we should design our experiments in such a way that they will not provide material to answer more than one question; it is often desirable to do several experiments at once, and in observational work the same set of data often provides information relevant to several different questions.† The rule refers to the analysis of the data, not to the data themselves.

Causality then survives only in the form that when a systematic variation is suggested in a sufficiently specific form to be tested we shall be willing to accept it if a significance test indicates that it is probably present. In the usual form of "determinism" it is superficially precise, but in fact so vague as to be useless (Jeffreys 1937*a*, p. 245). Determinism may be true, but cannot be tested. Randomness may be true or untrue in any particular case, and can be stated and tested; and therefore it is from the extreme of randomness that we must start and not from that of determinism. To disprove it may take four observations or 100,000, and no rule can be laid down in advance as to how many it will take; this is expressed in the theory by admitting that a new parameter suggested may account for any fraction of the outstanding variation. Randomness, it must be noticed, is not a statement of fact but a working hypothesis used about the outstanding variation at any stage of the investigation. When it is adopted the joint probability of the observations is the product of their separate probabilities given specified value of the parameters, and this is the condition for the validity of Bernoulli's theorem. The theorem would however be unacceptable if at any stage a new parameter is found to be relevant, and in fact new parameters are usually found relevant sooner or later. Thus the assertion of Bernoulli's theorem into the indefinite future is equivalent to saying that we shall never know of more relevant parameters than we do now, and would be a denial of any possibility of progress. Any apparent instance of the theorem is a danger signal, indicating that too much has been assumed (Jeffreys 1936*b*, p. 343). Alleged experimental proofs of the theorem really show only that if any other parameters are present than those originally contemplated their influence is too small to be shown by the number of observations available. But randomness as a working hypothesis is valid simply because it leads to inferences capable of being tested and can be revised when need is shown.

---

* Ockham's principle, under some such name as the economy of hypotheses, is actually used in three different senses. (1) In Whitehead and Russell's "Principia Mathematica", the number of postulates of logic is reduced to a minimum, though many results shown to follow from the postulates would ordinarily be considered as intuitively obvious as the postulates. The motive is here apparently aesthetic. On the other hand the fact that a person holds two beliefs at the same time, though it affords some evidence for their consistency, is not decisive evidence. The proof that one follows from the other therefore eliminates some of the uncertainty associated with possible imperfections of the human mind. (2) Parameters in a scientific law that make no contribution to the results of any observation can be eliminated mathematically, leaving the observations to be described only in terms of the relevant parameters. When this is done an economy of statement is achieved, but no improvement in representing facts. (3) The third is the statement in the text, which refers to the number of parameters to be admitted as relevant. The second and third principles are often confused. The second is always a pure tautology, but it is often stated as "the principle of the rejection of unobservables" and used to deny the relevance of any parameter not yet considered. It then becomes an *a priori* statement that future observations *must* follow the laws already considered, whatever the observations themselves may say. Such an inference into the future must be an inductive inference based on probability, because it is logically possible that the observations may disagree with prediction. The third principle deals with such inferences; but the second when used in this way involves a logical fallacy.

† E.g. several sets of data given by Fisher (1936) test two Mendelian ratios and linkage; but he uses separate tests for all three. I have reworked a number of his examples by my methods but have not found a case yet where the actual decisions differ, though his criteria differ and differences would be expected to occur occasionally.

3—*The Revision of Prior Probabilities*—The object of this work has been to give formal expression to the principles best laid down, I think, in the "Grammar of Science". Pearson says (1911, p. 12): "Now this is the peculiarity of scientific method, that when once it has become a habit of mind, that mind converts all facts whatsoever into science. The field of science is unlimited: its material is endless, every group of natural phenomena, every phase of social life, every stage of past or present development is material for science. *The unity of all science consists alone in its method, not in its material.* The man who classifies facts of any kind whatever, who sees their mutual relation and describes their sequences, is applying the scientific method and is a man of science." The fundamental postulates of such a system must be absolutely general; any restriction on their applicability is an assertion that we are entitled to use one criterion of validity in one subject and a different one in another. My object in this series of papers has been to find a set of simple and general postulates that will make it possible to assimilate any set of observed data into the general mass of scientific knowledge; and they have led to satisfactory rules for estimation and for testing new parameters in the fundamental problems of sampling, contingency, function-fitting, and the comparison of means of measured values.

It must be noticed that all my assessments of prior probabilities are statements of previous ignorance. They would be applicable in a new subject, but not necessarily in one that is already highly developed. It is a fact that our degrees of confidence in an inference are often altered when there is a change in the data; the whole object of the theory is to give a formal account of how this can take place. The function of the assessments is to show how it is possible to begin. They cannot be expected to hold indefinitely at all stages of knowledge, because if they did they would lead to illegitimate inferences about frequency in the long run by Bernoulli's theorem. It is a legitimate question, therefore, to ask what assessments should replace them in any advanced subject, allowing for previous experience in that subject. Pearson has noticed this point (1907, p. 366). "I start as most mathematical writers have done, with the equal distribution of ignorance, or I assume the truth of Bayes's theorem. I hold this theorem not as rigidly demonstrated, but I think with Edgeworth that the hypothesis of the equal distribution of ignorance is, within the limits of practical life, justified by our experience of statistical ratios, which *a priori* are unknown, i.e. such ratios do not tend to cluster round any particular value. Chances lie between 0 and 1, but our experience does not indicate any tendency of actual chances to cluster round any particular value in this range. The ultimate basis of the theory of statistics is thus not mathematical but observational. Those who do not accept the hypothesis of the equal distribution of ignorance are compelled to produce definite evidence of the clustering of chances, or to drop all application of past experience to the judgment of probable future statistical ratios. It is perfectly easy to form new statistical algebras with other clustering of chances." Further comment on this passage seems needed. I should say that if at the outset we are ignorant of a chance, Bayes's assessment of the prior probability is merely the formal way of saying so. It has nothing to do with observation, but simply translates from ordinary language into probability notation. The postulate that probabilities are orderable, and the addition rule for disjunctions, have already been given, and granting these there is no other way of saying "I do not know the chance". A sceptic may say that the statement "on data r, proposition p is more probable than q" is meaningless, that is, that prediction is meaningless. The position would be logically consistent; I only say that I know nobody that holds it consistently (1936*c*, p. 324). On the other hand Bayes's assessment is a statement of initial ignorance for one class and only one. But when we try to estimate other chances the questions would at once arise "Are all chances equal?" which a significance test answers in the negative; and then "Do chances show any significant departure from a uniform distribution?" Edgeworth's statement is then equivalent to saying that no significant departure from uniformity has been found. If it is correct we are still entitled to proceed, saying that any chance is equally likely in a new class, and at this stage a few further instances would alter the assessment for the next less than it would at an earlier stage. But actually the statement is incomplete. Pearson recognizes this elsewhere (1911, p. 141) when he comments on the tendency of sampling ratios to cluster at 0 and 1, and proposes a modification of Laplace's theory to take this into account. Again, no geneticist would agree that chances in his subject are uniformly distributed. The results given earlier in this paper would apply to the conditions at the time of Mendel's original experiment, but a modern Mendelian would be entitled to use the probabilities indicated by the observed frequencies of 1 : 1, 1 : 3, 3 : 5, 0 : 1; ..., ratios in interpreting a new experiment, and in fact does so roughly. Mendel's first results each rested on about 8000 observations; some hundreds would now usually be considered enough, and this corresponds to the fact that all that is now needed is to establish a high probability for one ratio compatible with the Mendelian theory against the others that have previously occurred and a background of other ratios attributable to differences of viability. A chemist wanting the molecular weight of a new compound would not content himself with a statement of his own determination. He carries out a complete analysis, finds one constitution consistent with all the data, and if he makes any further use of the results he will adopt a molecular weight calculated from the "International Tables". The uncertainty will be that of the calculated value, not that of his own. Thus previous information is habitually used and allowed for, whether explicitly or not, and it is not in all subjects that the previous information is of the type considered by Pearson in the passage quoted.

On similar grounds I should accept Dirac's argument (1937) on the clustering of non-dimensional physical parameters near powers of 10^{39} as a basis for inductive inference.

Differences in procedure in different subjects are therefore explained by differences in the nature of previous results, allowed for in a way equivalent to a reassessment of the prior probability based on observation. There is no need to assume any difference in fundamental principles, which themselves provide means of making such reassessments. It is in fact desirable that the results of a subject should be analysed at convenient intervals to see whether any alteration will be needed for future use, in order that the inferences shall represent as closely as possible the knowledge available at the times when they are made. Any subject in its development obtains the kind of information that is needed to bring its prior probabilities up to date. At present, however, we must be content with approximations, and in some subjects at any rate there appears to be no need for any immediate modification of the assessments used to express ignorance. Fisher's fiducial limits, which have been extensively used, differ a little from my critical values for significance, but the differences are slight for ordinary numbers of observations and the decisions would rarely be different. The fact that Fisher's rules have led to considerable progress where they have been applied, without any serious need for revision having been found, seems to indicate that in these subjects my assessments of the prior probabilities do not require much alteration to bring them into accordance with present knowledge. Thus if Fisher had found that only 1/20 of the differences found statistically lay beyond his 5 % limit he would presumably have noticed it and inferred that the limit for significance should be drawn at a higher level to reduce the risk of accepting random variations as significant. In subjects where statistical methods have hitherto had little application the assessments are applicable as they stand. It is clear that we cannot revise them in the same way in all subjects; experience in genetics is relevant to other problems in genetics, but not in earthquake statistics.

Attention is called to the fact that in my tests the ratio of the critical value of a difference to the standard error of the latter varies a little with the number of observations. A difference of twice the standard error may be just significant when it rests on five observations, but not when it rests on 100. For application of the tests it is therefore necessary to know the number of observations, and in many cases this is not given explicitly in published work and can be disentangled with great difficulty, if at all. In other words a difference of $1{\cdot}0 \pm 0{\cdot}5$ units may be worth considering further if it rests on five observations each with a standard error of $1{\cdot}2$ units; if it rests on 100 observations each with a standard error of 5 units it is not. This comes from pure probability theory and does not allow for the possibility of systematic error of observation, which might be considered at a later stage and would accentuate the effect.

SUMMARY

A mistake in two significance tests previously given by the author is corrected; the difference between these appeared to be much more than could be accounted for by the difference in the experimental conditions

considered, and is now almost entirely removed. The revision leads to a restatement of Ockham's principle for practical application, and to a discussion of the revision of prior probabilities as knowledge in a subject advances.

REFERENCES

Dirac 1937 *Nature, Lond.*, **139**, 1002.

Fisher 1936 "Statistical Methods for Research Workers." Edinburgh: Oliver and Boyd.

Jeffreys 1935 *Proc. Camb. Phil. Soc.* **31**, 203–22.

— 1936*a* *Proc. Camb. Phil. Soc.* **32**, 416–45.

— 1936*b* *Phil. Mag.* **22**, 337–59.

— 1936*c* *Mind*, **45**, 324–33.

— 1937*a* "Scientific Inference" (reissue). Camb. Univ. Press.

— 1937*b* *Proc. Camb. Phil. Soc.* **33**, 35–40.

Keynes 1921 "Treatise on Probability." London: Macmillan.

Pearson, K. 1907 *Phil. Mag.* **13**, 365–78.

— 1911 "The Grammar of Science." A. and C. Black.

Researches on the Chemistry of Coal

VIII—Comparative Studies of the Macro-Constituents of Bituminous Coals

BY WILLIAM A. BONE, D.SC., F.R.S. AND BASIL J. A. BARD, PH.D.

(*Received* 7 *July* 1937)

Often during recent discussions upon the chemical constitution of coal it has been asked whether or not there is any recognizable difference between the essential chemical structures of the principal macro-constituents (i.e. the "bright", "dull", and "fusain") of bituminous coals. Accordingly the authors have carried out comparative studies of the benzene-pressure extractions and alkaline permanganate oxidations of such constituents with a view to answering the question, and the present paper summarizes their findings.

At the outset it should be noted that, however carefully its macro-constituents may have been separated out from a given coal, it is improbable that any one of them can be regarded as a simple chemical substance, for so far no evidence has been forthcoming of their chemical homogeneity.

Fischer, Peters and Cremer (1932) carried out benzene-pressure extractions of the "bright" and "dull" constituents of certain German coals without finding any material difference between the yields of the various fractions of the crude extract, although the "bright" yielded rather more total extract than the "dull" coal.

This result was subsequently confirmed by Bates (1933) who had subjected certain representative British coals to similar extractions. But while agreeing that the "bright" yielded a somewhat larger total extract than the corresponding "dull" coal, he considered that the corresponding fractions into which the primary extracts from such coals can be divided by suitable successive solvent treatments have not always the same "individualities"; albeit he was somewhat vague as to the differences between them.

In this connexion it may be recalled that when Francis and Wheeler (1925) made ultimate analyses of the "ulmic acids" which they had derived from the "vitrain" and "durain" (i.e. "bright" and "dull") portions of a certain bituminous coal, their results, namely—

	Percentages of		
Ulmic acid from	C	H	O
"Vitrain"	64·6	2·6	32·8
"Durain"	65·4	3·0	31·6

—scarcely suggested any fundamental chemical difference between them.

EXPERIMENTAL

1—*The Coals*—The following three typical British coals were selected for the present investigation:

(i) *Busty "A"*, a well-matured Durham hard caking coal upon which a considerable amount of work has been done earlier in these investigations. The particular sample used in the present experiments had been stored for some years in the laboratory during which it had become slightly oxidized, its oxygen content having increased by about 1 % in 6 years.

(ii) *Dalton Main*, a medium caking coal from the Parkgate seam at Rotherham, with "bright" and "dull" constituents in the approximate ratio 3 : 2.

(iii) *Mitchell Main*, a strongly caking coal of the para bituminous series from the Parkgate seam at Barnsley, with "bright" and "dull" constituents in the approximate ratio 3 : 1.

Of these (ii) and (iii) were selected because of their having been extensively used by the Fuel Research Board, to whom we were indebted for the samples employed in our experiments. As received, they had already been roughly separated into "bright" and "dull" portions, and it only remained for us to complete the process.

2—*Separation of their Macro-Constituents*—First of all each coal was broken down into small lumps which were then individually scrutinized, any "mineral charcoal" (i.e. "fusain") being scratched out with a spatula and put aside for subsequent investigation. From the remainder were rejected all pieces containing both "dull" and "bright" constituents so intimately associated as to appear of a texture intermediate between the surrounding homogeneous "bright" and "dull" constituents which were afterwards hand-separated. Each was subsequently ground and sieved until it passed through a 20 but was retained by a 40-mesh sieve, such being the degree of fineness required for the benzene pressure extractions. Another part of each was further ground so as to pass through a 100-mesh sieve, such being used for the alkaline permanganate oxidations. The "fusain" was also reduced to similar degrees of fineness before use.

3—*Proximate and Ultimate Analyses of the Separated Constituents*—From the data shown in Table I it will be observed that in each case the "bright" contained notably less mineral matter and yielded more "volatiles" at 900° than the "dull" coal, while the "fusain" yielded the least "volatiles" of all. Also the carbon-contents were directly but the hydrogen- and oxygen-contents inversely in the order "bright"—"dull"—"fusain", a circumstance which was strongly reflected in the C/H ratios.

4—*Benzene Pressure Extractions*—These were carried out on the "bright" and "dull" constituents of the Busty and Dalton Main coals, respectively,

TABLE I—PROXIMATE AND ULTIMATE COMPOSITIONS OF THE SEPARATED CONSTITUENTS (DRY)

	Proximate analysis of dry coal			Ultimate composition of dry ashless coal					Ratio C/H
	% ash	% volatiles at 900°	Residue	C	H	N	S	O (diff.)	
Busty "A"									
Bright	2·20	24·65	Strong grey coke	85·85	4·60	1·45	1·10	7·00	18·6
Dull	4·65	21·80	Black weak coke	86·90	4·65	1·25	0·85	6·35	18·7
Fusain	3·80	12·15	Black powder	89·60	3·50	0·75	0·35	5·80	25·7
Dalton Main									
Bright	1·30	33·70	Hard silver-grey coke	84·70	5·70	1·75	1·10	6·75	14·9
Dull	5·15	32·30		86·45	5·20	1·45	1·05	5·85	16·6
Fusain	10·90	13·45	Black powder	88·90	4·10	0·50	0·55	5·95	21·7
Mitchell Main									
Bright	1·90	34·45	Hard silver-grey coke	86·30	5·80	1·60	0·95	5·35	16·3
Dull	3·80	27·10		88·85	5·20	1·55	0·70	3·70	23·7
Fusain	6·95	11·65	Black powder	92·30	3·45	0·70	0·65	2·90	31·7

and in each case the primary extract was subsequently fractionated according to the procedure already given by Bone, Pearson and Quarendon (1924). The results, as shown in Tables II, III, and IV, do not suggest any material difference between either the yields or the ultimate compositions of the corresponding fractions from the "bright" and "dull" constituents, respectively, of the same coal.

TABLE II—BENZENE PRESSURE EXTRACTIONS. YIELDS OF FRACTIONS

	Percentage yields of fractions on D.A. coal					Total extract
	I*a*	I*b*	II	III	IV	%
Busty						
Bright	0·5	1·9	2·9	0·4	6·8	12·5
Dull	0·3	2·5	3·0	0·3	6·3	12·4
Dalton						
Bright	0·5	2·4	1·4	1·4	8·8	14·5
Dull	0·5	1·6	2·5	0·8	7·6	13·0

TABLE III—ULTIMATE COMPOSITIONS OF FRACTIONS FROM BUSTY COAL

Fraction	I*b*		II		III		IV	
%	Bright	Dull	Bright	Dull	Bright	Dull	Bright	Dull
C	87·2	86·6	90·5	88·7	72·3	75·7	85·4	87·9
H	7·9	10·2	6·2	7·4	6·2	5·9	5·5	5·2
N	0·6	0·5	1·2	1·4	—	—	1·4	1·2
S	Trace	Trace	1·0	1·0	Trace	Trace	0·6	0·6
O (diff.)	4·3	1·7	1·1	1·5	—	—	7·1	5·1

TABLE IV—ULTIMATE COMPOSITIONS OF FRACTIONS FROM DALTON MAIN COAL

Fraction	I*a*		I*b*		II		III		IV	
%	Bright	Dull	Bright	Dull	Bright	Dull	Bright	Dull	Bright	Dull
C	85·8	85·8	88·1	88·5	87·8	87·2	72·4	76·3	85·6	86·2
H	9·8	10·8	8·5	7·5	6·7	7·2	6·3	6·2	6·4	5·7
N	0·1	0·1	0·4	0·4	0·8	1·0	1·8	1·6	1·6	1·2
S	Nil	Nil	Trace	Trace	0·7	0·8	Trace	Trace	0·5	0·5
O (diff.)	4·3	3·3	3·0	3·6	4·0	3·8	19·4	15·8	5·9	6·4

5—*Alkaline Permanganate Oxidations*—The "bright" and "dull" constituents of all three of the selected coals were separately oxidized under like conditions by means of boiling alkaline potassium permanganate, according to the procedure already described by Bone and Quarendon (1926), and "carbon balances" for the yields of carbon dioxide, acetic, oxalic and benzene carboxylic acids were subsequently drawn up. For if there is any major difference between the chemical structures of the "bright" and "dull" constituents of a given coal it should be reflected in the carbon distribution between their various oxidation products. But, as will be seen from the results detailed in Table VI, while there was never any qualitative difference between the oxidation products obtained from the "bright" and "dull" constituents, respectively, the "dull" always yielded a somewhat greater proportion of its carbon as benzene carboxylic acids and a correspondingly lesser proportion as oxalic acid than the "bright" constituent, a circumstance which may be regarded as indicative of some slight difference between their fundamental "benzenoid" structures. And in this connexion it should be noted that the "dull" was always rather more resistant to oxidation than the "bright" constituent.

This being so, it seemed of interest to make a comparative study of the alkaline permanganate oxidation of the "fusain" or "mineral charcoal" constituent, and for this purpose that of the Busty and Dalton Main coals was separately examined. In each case it proved to be very much more resistant to oxidation than either of the corresponding "bright" or "dull" coals; for even after 200 hr. only about half the original "fusain" had been oxidized, the remainder being recovered practically unchanged. The "carbon balances" obtained in each case are given in Table V.

TABLE V

	Fusain from	
Percentage of original carbon	Busty	Dalton
Unchanged	50·9	50·2
Yielded as CO_2	21·9	23·2
Yielded as CH_3COOH	0·7	0·5
Yielded as oxalic acid	5·0	3·8
Yielded as benzene carboxylic acids	20·1	20·8
Total accounted for	98·6	98·5
Time of oxidation (hr.)	192	160

Referred to the portion of the "fusain" actually oxidized in each case, the "carbon balances" would be:

Of "fusain" oxidized	Busty	Dalton
Percentage of original C obtained as:		
CO_2	45·3	47·4
CH_3COOH	1·4	1·0
Oxalic	10·3	7·7
Benzene carboxylic acids	41·6	42·4
Total	98·6	98·5

On comparing these results with those yielded by the corresponding "bright" and "dull" *constituent* in each case (Table VI) it will be seen that while the "fusain" yielded substantially more carbon dioxide than either of the other two, in the case of "Busty" its yield of benzenoid acids was substantially less than, but in the case of "Dalton" much the same as, that given by either of the other two constituents, a circumstance which again points to some slight difference between their fundamental "benzenoid" structure.

TABLE VI—RESULTS OF ALKALINE PERMANGANATE OXIDATIONS OF BRIGHT (b) AND DULL (d) COALS

Coal ...	Busty		Dalton Main		Mitchell Main	
	Bright b	Dull d	*Bright* b	Dull d	*Bright* b	Dull d
Percentage of original carbon obtained as:						
CO_2	35·9	38·8	40·0	39·2	40·0	39·2
CH_3COOH	1·6	1·4	1·9	2·0	1·7	1·6
$\begin{array}{c}COOH\\ \cdot\\ COOH\end{array}$	12·9	10·9	15·1	11·9	15·9	12·1
Benzene carboxylic acids	46·7	47·8	39·9	45·1	41·0	46·0
Total	97·1	98·9	96·9	98·2	98·6	98·9
Time taken (hr.)	26	172	21	29	45	50
Ratio $\frac{KMnO_4}{Coal}$ used	9·3	10·0	9·7	9·9	11·0	10·7

It may be added that, in addition to the foregoing oxidations under "carbon-balance" conditions, bulk oxidations were undertaken, using 50 g. of the particular constituent in each case, and afterwards converting the mixture of benzenoid acids obtained into methyl esters which were finally fractionally distilled under reduced pressure (1 mm.) into those passing over (*a*) up to 140°, (*b*) between 140° and 190°, and (*c*) residue. No marked difference between their compositions was disclosed thereby.

Our thanks are due to the Fuel Research Board for providing us with representative samples of Dalton and Mitchell Main coals employed in the investigation.

SUMMARY AND DISCUSSION

(i) The "bright", "dull", and "fusain" constituents of Busty "A" Dalton Main and Mitchell Main coals, respectively, have been separated and each of them has been separately submitted to (*a*) proximate and ultimate analyses, (*b*) benzene pressure extraction with subsequent fractionation of the primary extract, and (*c*) oxidation by means of boiling alkaline permanganate, in accordance with the usual standard procedure in each case.

(ii) In regard to their proximate and ultimate compositions, while, with the exception of Busty "A", both their mineral and carbon contents were in the ascending order "bright", → "dull", → "fusain", the reverse was always the case with their yield of "volatiles" at 900° and their hydrogen, nitrogen and sulphur contents.

(iii) In regard to the results of the benzene pressure extractions of the "bright" and "dull" constituents, respectively, of the same coal, no material difference was observed between either the yields or ultimate compositions of the corresponding fractions of the primary extract in each case.

(iv) In regard to the alkaline permanganate oxidations of the three macro-constituents of each coal, while their resistance to oxidation markedly increased in the order "bright" → "dull" → "fusain", no great difference was observable either qualitatively, as between the resulting products, or quantitatively, as regards the distributions of the original carbon between them, save only that the percentage yielded as carbon dioxide by "fusain" was always materially greater than that yielded by either of the two constituents.

(v) Viewed as a whole, there is nothing in the evidence suggestive of more than quite minor differences between the fundamental benzenoid structures of the three principal macro-constituents of a given coal, although the "dull" appears to be rather more matured than the corresponding "bright" coal.

REFERENCES

Bates, W. E. 1933 *Tech. Pap. Fuel Res., Lond.*, No. 37.
Bone, W. A., Pearson, A. R. and Quarendon, R. 1924 *Proc. Roy. Soc.* A, **105**, 608.
Bone, W. A. and Quarendon, R. 1926 *Proc. Roy. Soc.* A, **110**, 537.
Fischer, F., Peters, K. and Cremer, W. 1932 *BrennstChem.* **13**, 364.
Francis, W. and Wheeler, R. V. 1925 *J. Chem. Soc.* **127**, 2237.

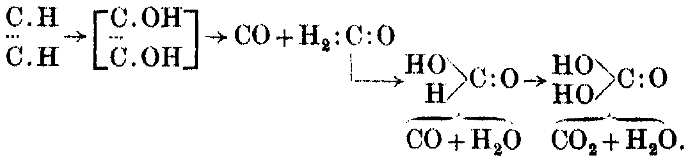
The Slow Combustion of Acetylene

By William A. Bone, D.Sc., F.R.S.
and J. E. Carruthers, B.Sc., Ph.D. (*Beit Research Fellow*)

(*Received* 7 *July* 1937)

In 1905 one of us in conjunction with Andrew published the results of extensive experiments upon the combustion of acetylene which showed that when a $2C_2H_2 + O_2$ or $C_2H_2 + O_2$ mixture was sealed up in borosilicate glass bulbs at atmospheric pressure and afterwards heated, reaction set in at 250°, or even at a somewhat lower temperature, and proceeded rapidly at 300°. Explosive combustion set in at about 350°, the ignition point being raised either by reducing the initial pressure or by addition of oxygen to the equimolecular mixture.

With regard to the slow combustion, it was shown that carbonic oxide and formaldehyde arise simultaneously at an early stage of the process, probably as the result of the thermal decomposition of an unstable $C_2H_2O_2$ which might possibly be $\begin{array}{c}\text{C.OH}\\ \vdots\\ \text{C.OH}\end{array}$, although only a form of polyglycolide $(C_2H_2O_2)_x$ was actually isolated. The formation of formaldehyde preceded that of steam, and the whole process, it was thought, might be represented by the scheme:



$$\begin{array}{c}\text{C.H}\\ \vdots\\ \text{C.H}\end{array} \rightarrow \left[\begin{array}{c}\text{C.OH}\\ \vdots\\ \text{C.OH}\end{array}\right] \rightarrow CO + H_2{:}C{:}O \rightarrow \underbrace{\begin{array}{l}HO\\ H\end{array}\!\!>C{:}O}_{CO + H_2O} \rightarrow \underbrace{\begin{array}{l}HO\\ HO\end{array}\!\!>C{:}O}_{CO_2 + H_2O}.$$

In 1930 Kistiakovsky in conjunction with Lenher and Spence used a flow method for investigating the reaction, mixtures of the two reactants being drawn at varying speeds through a tube maintained at some temperature between 280 and 322°, and condensible products removed immediately on leaving the furnace by means of U-tubes cooled in a solid CO_2-ether mixture. A short "inductive period" of about 30–40 sec. was observed, and packing the reaction tube with pieces of broken glass led to a marked reduction in the reaction rate, pointing to the reaction being predominantly homogeneous and of the "chain" type.

Subsequently Spence proposed a complicated "chain" mechanism involving glyoxal molecules in two different states of activation as the "chain carriers". But while Kistiakowsky and Lenher found that in the presence of acetylene formaldehyde is further oxidized to formic acid, CO_2 and H_2O at a rate some hundred times faster than when alone, Stacie and Macdonald recently observed that the addition of formaldehyde to a $C_2H_2+2O_2$ mixture definitely retarded its interaction, and that the addition of glyoxal had no accelerating effect.

In view of such seemingly contradictory observations, and of improvements in experimental methods during more than thirty years that had elapsed since the work of Bone and Andrew, the authors decided to reinvestigate the subject, employing the procedure which had been applied so successfully in the research laboratories of the Imperial College, London, to the slow combustions of methane, ethane and ethylene, and the present paper gives an account of their results.

EXPERIMENTAL

1—*General Procedure*

P_2O_5-dried mixtures of highly purified acetylene and oxygen were admitted into an evacuated silica reaction vessel *A* (fig. 1) maintained at some selected constant temperature between 265 and 195° in an electric resistance

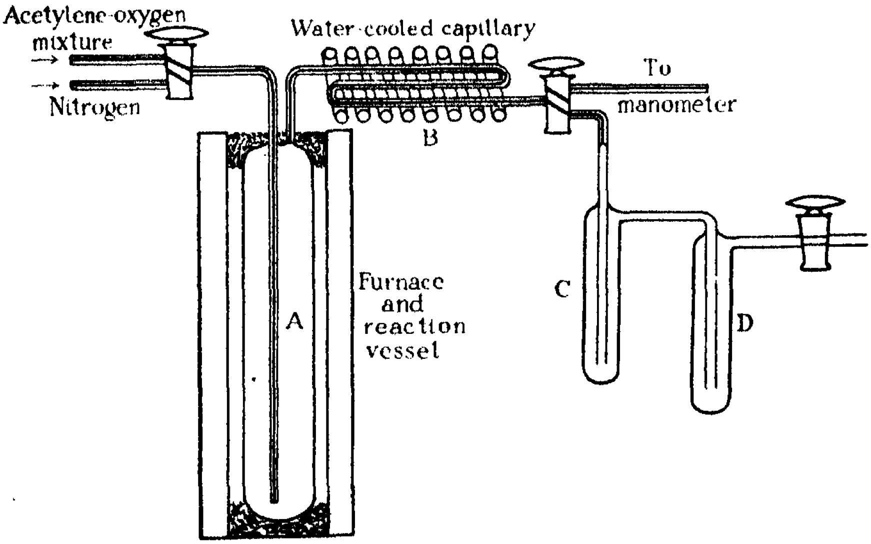


FIG. 1—Diagram of apparatus.

furnace, the reaction vessel being connected on the one hand through the water-cooled capillary tube *B* with a suitable manometer, and on the other with the well-cooled evacuated bulbs *C* and *D* whereby any condensable products were effectively removed before the gaseous ones were collected over mercury for subsequent analysis.

In most experiments the reaction was stopped suddenly at some predetermined point by removing the reaction vessel from the furnace and quickly plunging it into ice-cooled water, whereupon the gaseous products were withdrawn for analysis, after which the reaction vessel was washed out with a known volume of distilled water and the washings analysed quantitatively for condensable products.

By such procedure not only were pressure-time curves for the reacting media obtained, but the gaseous and condensable products were removed at any selected stage of the reaction and completely analysed; and from the combined results in each case carbon-hydrogen-oxygen balances were subsequently drawn up.

2—*Transition from Slow to Explosive Combustion*

Although reaction in a $2C_2H_2 + O_2$ mixture proceeded slowly at atmospheric pressure within the temperature range (265–295°) referred to, it exploded after about 50 sec. on being admitted to the reaction vessel at 300°, carbon being deposited and the gaseous products consisting almost entirely of carbonic oxide and hydrogen together with traces of carbon dioxide, methane and surviving reactants but without any formation of steam, in accordance with the empirical equation:

$$2C_2H_2 + O_2 = 1{\cdot}85C + 2{\cdot}0CO + 1{\cdot}8H_2 + 0{\cdot}01CO_2 \\ + (0{\cdot}05CH_4 + 0{\cdot}07C_2H_2 + 0{\cdot}05O_2).$$

On the other hand, the products of the slow combustion within the temperature range referred to were found to consist of oxides of carbon, steam, glyoxal, formaldehyde, and formic acid, together with traces only of hydrogen. No "peroxide" was ever detected in them.

The Pressure-Time Curves—Typical pressure-time curves obtained for $C_2H_2 + O_2$ at 265° and for $2C_2H_2 + O_2$, $3C_2H_2 + O_2$ and $4C_2H_2 + O_2$ mixtures, respectively, at 272° are shown in figs. 2–5. It will be seen that in the first case there was a slight pressure-fall of 2·5 mm. during the first 20 min. followed by a continuous rise of 19 mm. as the reaction proceeded to completion, the initial pressure having been 500 mm. With the $2C_2H_2 + O_2$ mixture at 272° there was a pressure-fall of 7·5 mm. during the first 12 min.

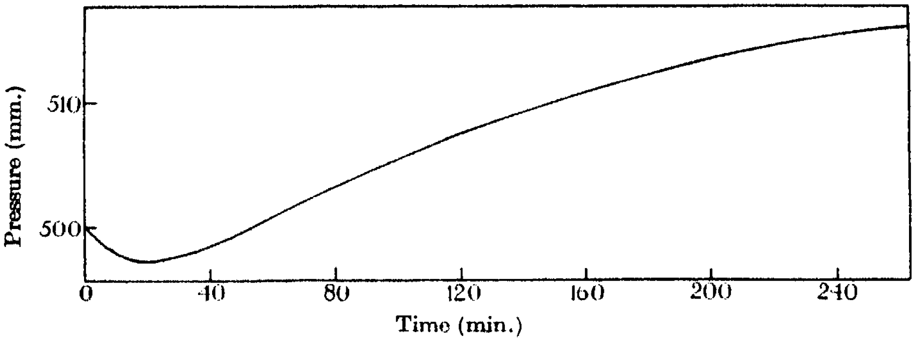


FIG. 2—$C_2H_2+O_2$. 265° C.

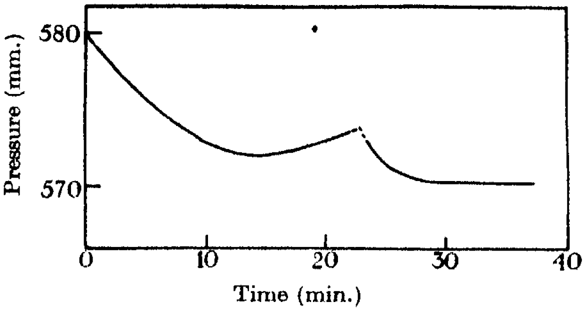


FIG. 3—$2C_2H_2+O_2$. 272° C.

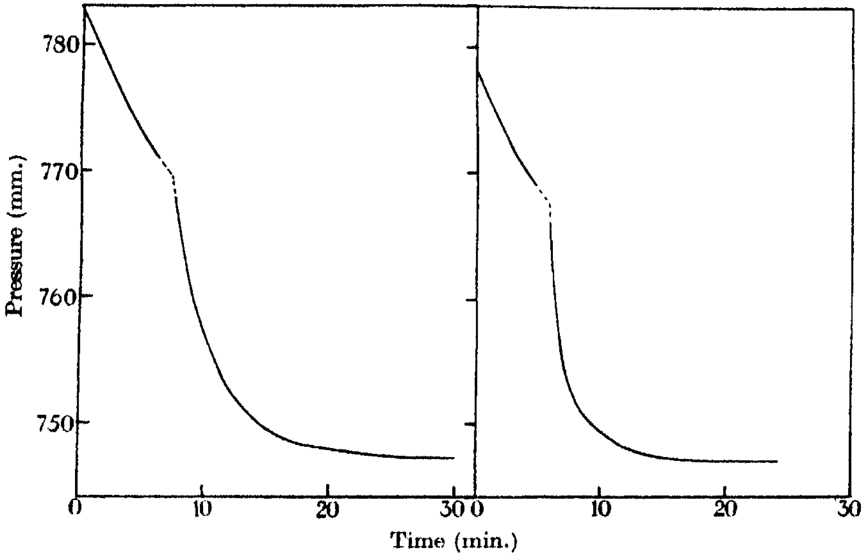
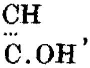


FIG. 4—$3C_2H_2+O_2$. 272° C. FIG. 5—$4C_2H_2+O_2$. 272° C.

followed by a slight rise during the next 10 min., after which there was an abrupt change, the pressure again falling somewhat to a final constant level.

With higher acetylene concentrations there was always a progressive pressure fall throughout an experiment with a marked abrupt change in the rate of fall after a definite time within the first 8 min. Such abrupt changes were marked by a rise of 0·7° in the temperature of the medium, doubtless due to some exothermic secondary reaction which became prominent at such juncture, and subsequent investigation showed that they occurred almost exactly at the point at which the whole of the original oxygen had disappeared.

It seems that whereas the secondary reaction referred to had been operative to some extent during the whole reaction, its effect upon the pressure-time curve only became apparent after all oxidation had ceased, when it became the dominant change in the system. And as its effect upon the pressure-time curve was only observable with media initially containing more than an equimolecular proportion of acetylene, it may have been due to the polymerization of some initially formed hydroxy-acetylene,  $\begin{matrix}\mathrm{CH}\\ \vdots\\ \mathrm{C.OH}\end{matrix}$, a supposition which was supported by the subsequent isolation of an as yet unidentified oily phenolic compound from the oxidation products, particularly from those of media initially containing high acetylene concentrations, as well as by other evidence obtained during the investigation.

3—*Variation of Reaction Rate with the Partial Pressures of the Reactants*

Taking the rates of oxygen disappearance as indicating the relative reaction velocities, the curves shown in fig. 6 summarize observations of such rates for $C_2H_2 + O_2$, $2C_2H_2 + O_2$ and $3C_2H_2 + O_2$ media, respectively, all initially containing the same oxygen concentration of 190 mm. and reacting at 272°. It will be seen that the fastest rate was given by the $3C_2H_2 + O_2$ medium.

Considering now the inverse of the time t required for half-reaction, we obtain from the curves:

Medium	$C_2H_2 + O_2$	$2C_2H_2 + O_2$	$3C_2H_2 + O_2$
$t_{\text{min.}}$	$73\frac{1}{2}$	$18\frac{1}{2}$	$8\frac{1}{2}$
$1/t$	0·0136	0·054	0·1175

The values of $1/t$ are in the ratio 1 : 3·97 : 8·7, or very nearly the same ratio as the squares of the corresponding acetylene concentrations, namely 1 : 4 : 9, in the three cases, proving the reaction velocity to have been proportional to the square of the acetylene concentration.

Similar curves for the rates of acetylene disappearance at 295° with $3{\cdot}8C_2H_2 + O_2$ and $1{\cdot}86C_2H_2 + O_2$ media having the same initial acetylene concentration were almost coincident, showing the reaction rate to have been independent of the oxygen concentration.

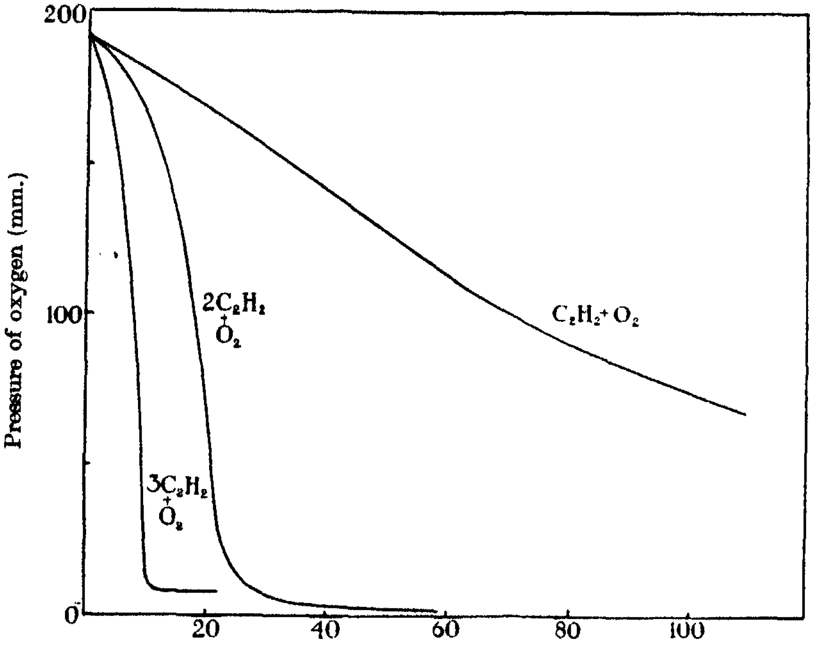


FIG. 6—Disappearance of oxygen at 272°.

4—*Compositions of Media initially* $2C_2H_2 + O_2$ *and* $C_2H_2 + O_2$, *respectively, at Various Stages during Reaction at* 265°

Several experiments were made with the object of tracing the rates of formation of the various products during the interactions of $2C_2H_2 + O_2$ and $C_2H_2 + O_2$ mixtures, respectively, at 265°; but for the purpose of illustration the result of only one of them in each case need be detailed. These are shown graphically in figs. 7 and 8, in which the partial pressures of the various components of the reacting media (acetylene, oxygen, glyoxal, formaldehyde, formic acid, CO, CO_2, and H_2O respectively) are plotted against time. It should be noted, however, that in order to include acetylene in the graph for the $2C_2H_2 + O_2$ medium (fig. 7), its curve had to be transformed to another scale by vertical downward displacement, as indicated in the right-hand margin thereof. No curve for H_2O formation is included in the graph for the $C_2H_2 + O_2$ medium (fig. 8).

Next after oxides of carbon and steam, formaldehyde was always the most prominent oxidation product, its concentration attaining a nearly constant maximum, in the case of the $2C_2H_2+O_2$ medium, after about four-fifths of the original oxygen had disappeared in about 40 min. With the $C_2H_2+O_2$ medium, however, its concentration, after attaining a maximum in about 60 min., fell off slightly toward the end of the experiment.

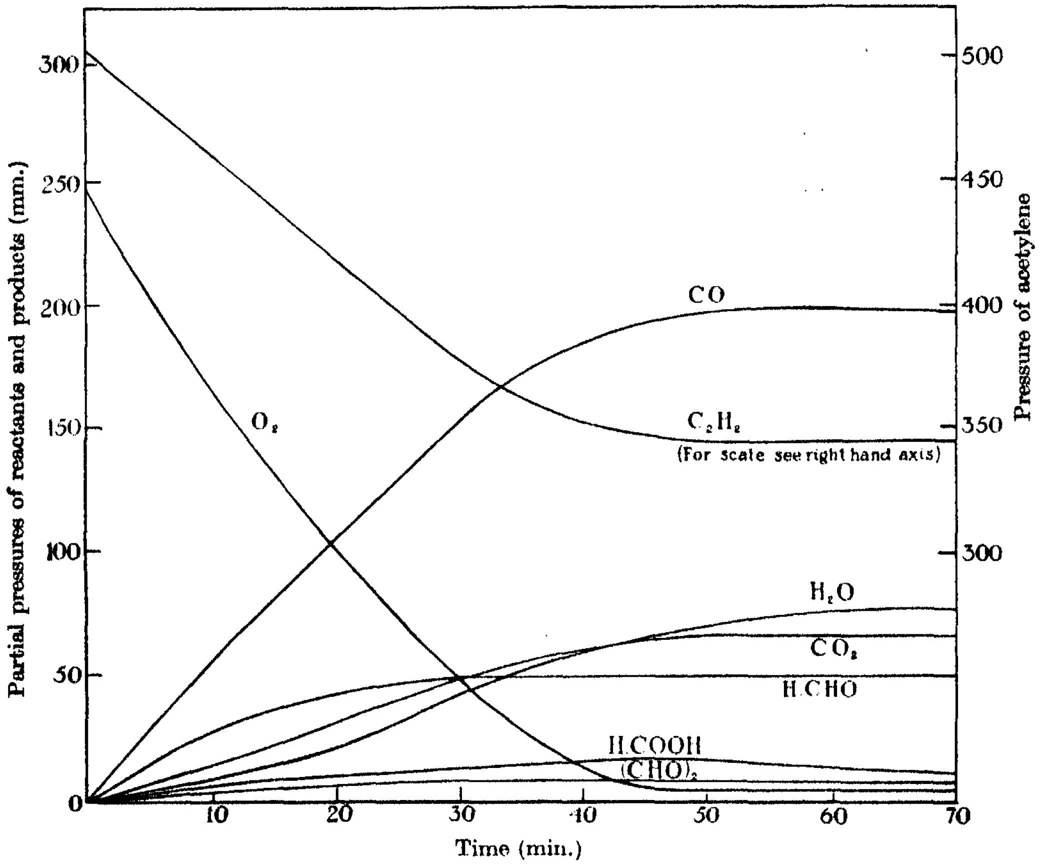


Fig. 7—Interaction of $2C_2H_2+O_2$ at 265° C.

5—*Carbon-Hydrogen-Oxygen Balances for the Reactions of* $2C_2H_2+O_2$ *and* $C_2H_2+O_2$ *Media, respectively, at 265°*

Attention is directed to the carbon-hydrogen-oxygen balances and other particulars shown in the accompanying Table I for the almost completed interactions of $2C_2H_2+O_2$ and $C_2H_2+O_2$ media, respectively, at 265°, from which it will be seen that practically the whole of the carbon,

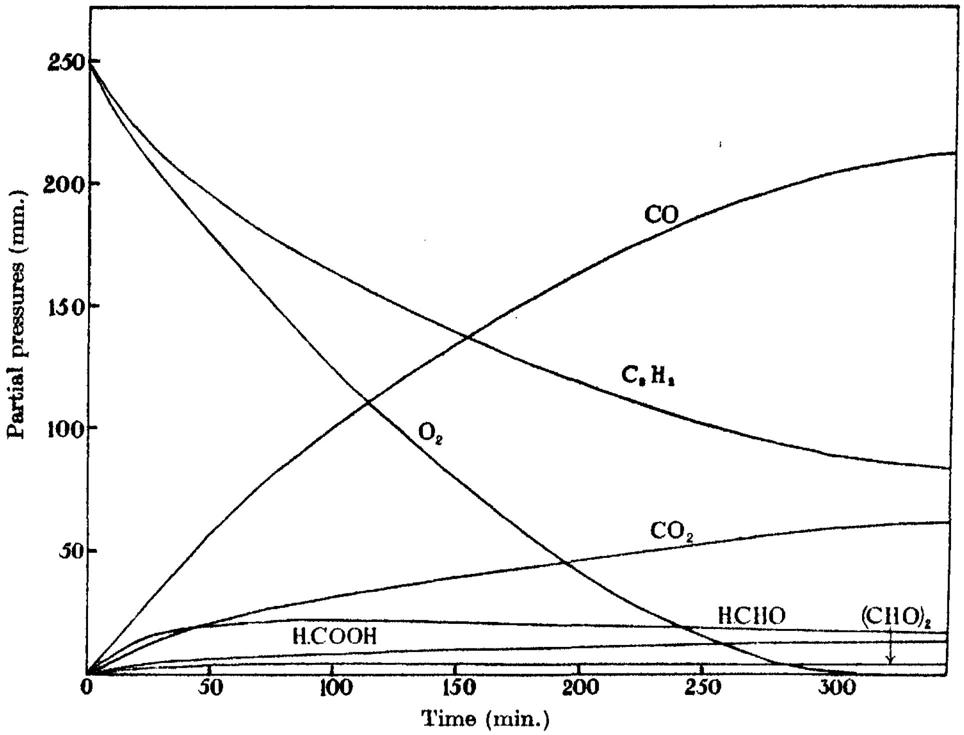


Fig. 8—Interaction of $C_2H_2 + O_2$ at 265° C.

TABLE I—REACTIONS OF $2C_2H_2 + O_2$ AND $C_2H_2 + O_2$ MEDIA AT 265°

Mixture		$2C_2H_2 + O_2$	$C_2H_2 + O_2$
Duration of reaction (min.)		41	298
Initial partial pressures at 265° (mm.)			
C_2H_2		497	245
O_2		242	247
		739	492
Final partial pressures at 265° (mm.)			
Gaseous	C_2H_2	354·0	88·0
	O_2	35·0	1·5
	CO_2	53·5	60·5
	CO	170·5	202·0
	H_2	Nil	3·5
Condensable	$(CHO)_2$	5·5	3·0
	CH_2O	49·0	17·0
	H·COOH	14·5	15·0
	H_2O	60·0	115·0
		742	505·5
C_2H_2 used (mm.)		143	157
O_2 used (mm.)		207	245·5
Ratio O_2/C_2H_2		1·45	1·56

Units in:	C	H_2	O_2	C	H_2	O_2
Original mixture	994·0	497	242	490	245·0	247
Final products	1006·5	483	248	476·5	241·5	247

hydrogen and oxygen originally present was accounted for in the end products enumerated. They also show how much more reactive was the $2C_2H_2+O_2$ than the $C_2H_2+O_2$ medium.

Summary and Discussion

1—The slow interactions of mixtures of acetylene and oxygen between the limits $4C_2H_2+O_2$ and $C_2H_2+O_2$ in a silica vessel at temperatures between 265 and 272° and atmospheric pressure have been experimentally studied as regards pressure-time curves, reaction rates, and the rates of formation of the various gaseous and condensable products.

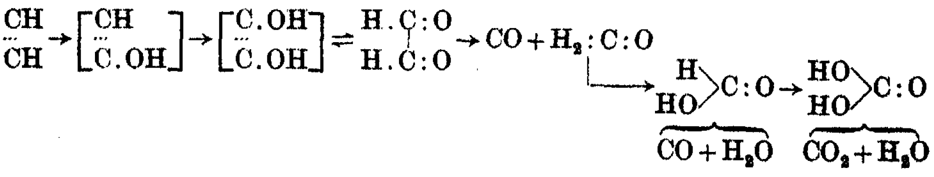
2—The pressure-time curves showed that (i) with $C_2H_2+O_2$ mixtures there was a slight pressure-fall during the first 20 min. followed by about an eight times greater rise as the reaction proceeded to completion, (ii) with $2C_2H_2+O_2$ mixtures, there was a slight pressure-fall during the first 12 min., followed by a short abrupt increase in the rate of fall down to a constant level, and (iii) with still higher initial acetylene concentrations, there was always a progressive pressure-fall throughout an experiment marked by an abrupt change in the rate of fall after a definite time period within 8 min. of the start.

3—The reaction velocity, as measured by the rate of oxygen-disappearance, was proportional to the square of the acetylene-concentration but independent of the oxygen-concentration.

4—The reaction products actually identified and measured were: glyoxal, formaldehyde, formic acid, oxides of carbon and steam; and in addition traces of an as yet unidentified oily phenolic substance were detected.

5—The evidence as a whole is consistent with the hydroxylation theory, assuming that the oily phenolic substance referred to may be regarded as a polymerization product involving unstable $\begin{matrix}C.H\\ \cdots\\ C.OH\end{matrix}$ molecules, and that the glyoxal arose by intermolecular change from an unstable $\begin{matrix}C.OH\\ \cdots\\ C.OH\end{matrix}$.

6—On such suppositions, the successive stages of the slow combustion may be represented as follows:



$$\begin{matrix}CH\\ \cdots\\ CH\end{matrix} \rightarrow \begin{bmatrix}CH\\ \cdots\\ C.OH\end{bmatrix} \rightarrow \begin{bmatrix}C.OH\\ \cdots\\ C.OH\end{bmatrix} \rightleftharpoons \begin{matrix}H.C{:}O\\ |\\ H.C{:}O\end{matrix} \rightarrow CO+H_2{:}C{:}O \rightarrow \underbrace{{}^{H}_{HO}\!>C{:}O}_{CO+H_2O} \rightarrow \underbrace{{}^{HO}_{HO}\!>C{:}O}_{CO_2+H_2O}$$

In conclusion the authors desire to thank the Governing Body of the Imperial College for the Beit Fellowship (1935–6) under which one of them carried out the experimental part of the investigation.

REFERENCES

Bone, W. A. and Andrew, G. W. 1905 *J. Chem. Soc.* **87**, 1232.
Kistiakovsky, Lenher and Spence 1930 *J. Amer. Chem. Soc.* **52**, 4837.
Spence 1932 *J. Chem. Soc.* p. 686.
Stacie and Macdonald 1936 *J. Chem. Phys.* **4**, 75.

The Diffusion of Pairs of Gases with Molecules of Equal Mass

By L. E. BOARDMAN, M.SC. AND N. E. WILD, M.SC.
Physics Department, University of Birmingham

(*Communicated by S. W. J. Smith, F.R.S.—Received* 1 *June* 1937)

INTRODUCTION

The problem of the self-diffusion of a gas has received considerable attention, and it is of special importance in the kinetic theory because of the similarity of the diffusing molecules. Chapman's exact classical kinetic theory (Chapman 1916, 1918) gives the expression:

$$D_{11} = \frac{k\eta}{\rho},$$

where D_{11} = coefficient of self-diffusion, η = coefficient of viscosity, ρ = density, k = a constant. The value of k varies from 1·504 to 1·200 for Maxwellian molecules (index of repulsive force 5) and rigid elastic spheres respectively. Massey and Mohr (1933), however, applying the quantum theory to collision processes in gases have deduced that for self-diffusion the factor given by the classical theory must be reduced by nearly $\frac{1}{2}$. From the point of view of this theory the molecules in true self-diffusion must be identical and indistinguishable,* so that the prediction in that case cannot be experimentally verified. Although the conditions of these experiments do not correspond to the case of self-diffusion of the quantum theory, it is a matter of considerable interest to examine the behaviour in diffusion of these pairs of gases which have in many respects a striking physical similarity.

---

* We are indebted to Dr. H. S. Massey for kindly giving us his opinion on this aspect of the question.

A method of calculating the coefficient of self-diffusion of a gas indirectly from a number of diffusion experiments was given by Kelvin (1904), but until recently no direct experimental measurements have been made. Harteck and Schmidt (1933) have attacked the problem by making diffusion experiments with normal hydrogen and a mixture rich in para-hydrogen.

In our investigation, diffusion experiments were carried out with the pairs of gases, nitrogen-carbon monoxide, and nitrous oxide-carbon dioxide. The members of these pairs of gases have been shown to have the same molecular weights, the same molecular diameters (Smith 1922), and the same molecular fields (Jeans 1921) and inter-molecular fields (Ibbs and Underwood 1927). Each pair consists therefore of two virtually identical gases, and the diffusion experiments on these pairs of gases are a close approximation to self-diffusion.

THEORY OF METHOD

The method consists in allowing the pure gases contained in vessels of equal length and equal uniform cross-section to diffuse into each other for a given time, after which the composition of the gas mixture in each vessel is determined.

If ρ represents the density of one gas at any point in the apparatus then the differential equation of diffusion is $\frac{\partial \rho}{\partial t} = D_{12} \frac{\partial^2 \rho}{\partial x^2}$ (*Handbuch der Physik*). In the case of self-diffusion the coefficient D_{12} is a constant and the equation is exact.

Solving this equation for the appropriate initial and boundary conditions gives the value of ρ for one gas at any point x along the diffusion tube (measured from the top) at any time t, whence by integration the masses o and u of the same gas in the upper and lower halves of the apparatus are obtained, giving finally

$$\frac{u-o}{u+o} = \frac{8}{\pi^2}\{e^{-D_{12}t(\pi/L)^2} + \tfrac{1}{9}e^{-9D_{12}t(\pi/L)^2} + \tfrac{1}{25}e^{-25D_{12}t(\pi/L)^2} + \ldots\},$$

where L is the total length of the diffusion apparatus (Obermayer 1880).

In the ratio $\frac{u-o}{u+o}$ the concentration of the gas may conveniently be substituted for the mass.

Using three terms of the above series Obermayer calculated the values of $\frac{u-o}{u+o}$ for various values of the argument $D_{12}t$, for his apparatus ($L = 86{\cdot}62$ cm.).

To use the table for a diffusion apparatus having a length L' the value of $D_{12}t$ given by Obermayer's table must be multiplied by $(L'/L)^2$ to give the correct value of $D_{12}t$.

EXPERIMENTAL DETAILS OF METHOD

The method outlined above was first devised by Loschmidt (1870) and used later by Obermayer (1880) and others (Schmidt 1904; Lonius 1909) for measuring the coefficient of diffusion of pairs of gases.

The arrangement of the diffusion tubes T_1, T_2 is shown in fig. 1. The tubes were made from copper tubing having a mean internal diameter of 1·3 cm. and walls nearly 3 mm. thick. The diffusion tubes were made as nearly as possible equal in length, and were closed by brass plugs P_1, P_2 soldered into the ends. Two circular brass disks D_1, D_2 pivoted at the centre were employed instead of a tap. The diffusion tubes were soldered into the disks at equal distances from the centre as shown in the diagram.

A pivot was screwed into the centre of the lower disk (fixed) about which the upper disk could be rotated by means of a "tommy" bar. A spring S, held in compression by a nut on the pivot, kept the faces of the disks always in contact. The contact surfaces were ground together, and made gastight with soft stop-cock grease.

The connecting tubes for filling and emptying the diffusion tubes were made of small-bore glass tubing, the metal-glass joints being made with picene wax. The tube C_1 connects the diffusion tubes with the rest of the apparatus (for filling the tubes and analysing the mixtures). C_2 leads to the atmosphere or may be connected to an automatic pump if it is desired to evacuate the apparatus without collecting the gases.

By a suitable rotation of the upper disk the upper diffusion tube may be (1) put into communication with the rest of the apparatus, (2) closed, or (3) brought into alignment with the lower diffusion tube. Exact alignment of the tubes was obtained by bringing scratch marks on the two disks into coincidence.

The whole apparatus was surrounded by a wooden case to shield it from draughts, and the experiments were carried out in a cellar which had no form of heating. Under these conditions the variation of temperature was very slow, and there was a constant temperature difference of 0·3° C. between the tubes, the upper tube having the higher temperature.

Length of diffusion apparatus = 89·09 cm. at 16° C.

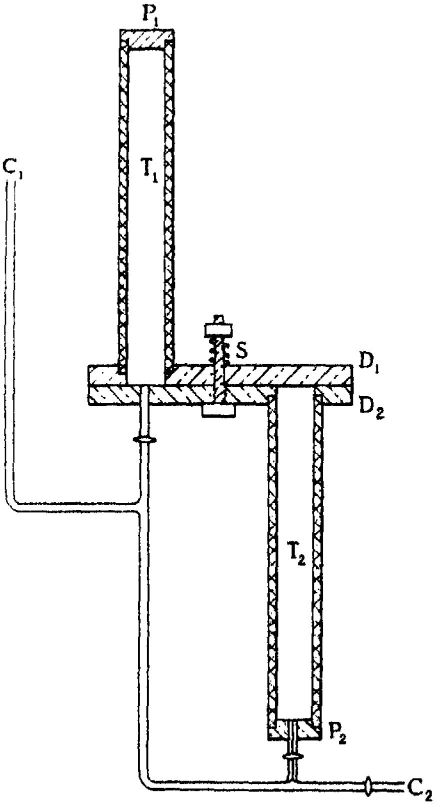


FIG. 1—The diffusion tubes.

All the gases used in the experiments were obtained from cylinders and stored in gas tubes over mercury. The diffusion tubes were evacuated first by means of an automatic pump and in the later stages by means of a Toepler pump. The diffusion tubes were filled from a gas burette at a pressure of 1 mm. of mercury above atmospheric pressure. After being allowed to stand for some time to reach a steady temperature the pressure in the diffusion tubes was made equal to atmospheric pressure (observed) by momentarily opening the tubes to the atmosphere. The diffusion experiment was then commenced, the time of duration being measured by a stopwatch.

At the end of the experiment the gas mixtures in the diffusion tubes were analysed.

The gases carbon monoxide and nitrogen have almost exactly the same thermal conductivity, and in order that the katharometer method (a method of analysis depending on a difference of thermal conductivity) could be used it was necessary to oxidize the carbon monoxide to carbon dioxide.

In the experiments with carbon monoxide and nitrogen, the gas mixtures were first left in contact with heated copper oxide until all the carbon monoxide had been oxidized to carbon dioxide, and then the mixtures of carbon dioxide and nitrogen were analysed by means of the katharometer (Daynes 1920).

In order to test the reliability of this method of analysing mixtures of nitrogen and carbon monoxide, pure carbon monoxide and a mixture of known proportions (calculated from partial pressures) were analysed by this method. The results of this test are shown in Table I.

TABLE I

% N_2 calculated from partial pressures	% N_2 analysed by katharometer
0·0	0·4
52·0	52·4

A volumetric analysis of the mixtures of carbon dioxide and nitrous oxide was made by absorbing the carbon dioxide by means of caustic potash.

Diffusion experiments were also carried out with the pairs of gases: hydrogen-carbon dioxide, hydrogen-nitrogen, and nitrogen-carbon dioxide, so that the coefficient of self-diffusion for nitrogen and carbon dioxide could be calculated by Kelvin's method. In these cases the gas mixtures could be analysed directly by the katharometer owing to the differences in the thermal conductivities of the gases used.

The purity of the gases used and the nature of the impurity are shown in Table II.

TABLE II

Gas	% purity	Impurity
CO_2	99·9	Air
CO	99·9	Air
N_2O	99·9	N_2
H_2	99·9	Air
N_2	99·7	O_2

EXPERIMENTAL RESULTS

The results of the experiments are shown in the following tables, the final values of the coefficients being reduced to 15° C. and 76 cm. pressure. For this purpose D_{12} was assumed to vary inversely as the pressure, and the temperature indices used were those determined experimentally by Obermayer's (1880) method. All the values of the coefficients of diffusion in the tables are in c.g.s. units.

TABLE III—NITROGEN-CARBON MONOXIDE.
PERIOD OF DIFFUSION = 15 MIN.

Exp. No.	Barometer cm. Hg	Mean temp. ° C.	Upper half % N_2	Lower half % N_2	D_{12}
1	74·6	17·7	18·7	82·5	0·219
2	73·9	16·3	18·3	82·7	0·211
3	75·2	15·2	82·8	17·5	0·206
4	74·5	16·0	17·9	82·8	0·208
					Mean = 0·211

Exps. 1, 2 and 4 were performed with nitrogen initially in the lower half of the apparatus, while Exp. 3 was performed with nitrogen initially in the upper half. In order to find if D_{12} was affected by a change in the period of diffusion an experiment was carried out in which the gases were allowed to diffuse for twice the period given in Table III; the result showed no appreciable change in D_{12}. The results of this experiment were as follows:

Barometer cm. Hg	Mean temp. ° C.	Upper half % N_2	Lower half % N_2	D_{12}
74·5	15·9	25·0	75·3	0·209

In this experiment nitrogen was initially contained in the lower half of the apparatus.

TABLE IV—NITROUS OXIDE-CARBON DIOXIDE.
PERIOD OF DIFFUSION = 60 MIN.

Exp. No.	Barometer cm. Hg	Mean temp. ° C.	Upper half % N_2O	Lower half % N_2O	D_{12}
1	74·1	13·8	75·3	25·5	0·107
2	76·2	12·4	24·9	75·6	0·108
3	74·5	12·9	25·0	75·3	0·106
4	73·3	14·1	25·5	75·0	0·107
5	73·2	14·3	75·0	25·0	0·105
6	75·3	14·1	24·9	75·4	0·106
					Mean = 0·107

Exps. 1 and 5 were carried out with nitrous oxide initially in the upper half of the apparatus, and Exps. 2, 3, 4 and 6 with nitrous oxide initially in the lower half of the apparatus.

TABLE V—HYDROGEN-NITROGEN. PERIOD OF DIFFUSION = 10 MIN.

Barometer cm. Hg	Mean temp. °C.	Upper half % H_2	Lower half % H_2	D_{12}
72·9	13·9	73·0	27·2	0·740
74·2	13·9	73·1	26·8	0·740
73·5	14·3	72·9	27·1	0·745
74·1	14·1	73·2	26·9	0·737
75·3	13·7	72·9	26·5	0·749
76·3	14·2	73·5	26·6	0·743
76·1	14·2	73·5	26·7	0·742
75·9	14·3	73·3	26·7	0·744
75·8	14·2	73·1	26·6	0·747
				Mean = 0·743

TABLE VI—HYDROGEN-CARBON DIOXIDE.
PERIOD OF DIFFUSION = 15 MIN.

Barometer cm. Hg	Mean temp. °C.	Upper half % H_2	Lower half % H_2	D_{12}
74·2	15·7	70·0	30·1	0·616
73·7	15·8	69·8	30·4	0·622
74·8	15·9	70·1	30·3	0·622
74·2	16·2	69·5	30·1	0·625
73·6	15·9	70·3	30·5	0·612
74·2	15·3	70·3	30·4	0·617
73·8	15·2	70·0	30·2	0·616
				Mean = 0·619

TABLE VII—NITROGEN-CARBON DIOXIDE.
PERIOD OF DIFFUSION = 40 MIN.

Barometer cm. Hg	Mean temp. °C.	Upper half % N_2	Lower half % N_2	D_{12}
75·6	16·1	75·2	24·8	0·158
75·5	16·9	75·0	25·0	0·159
75·6	15·7	75·2	24·8	0·158
75·7	16·3	75·2	24·9	0·159
73·7	16·1	74·6	24·6	0·156
74·7	15·3	75·1	24·4	0·155
				Mean = 0·158

DISCUSSION OF RESULTS

The mean value of D_{12} for nitrogen and carbon monoxide, which is to be regarded as the classical coefficient of self-diffusion D_{11} of either gas, was found to be 0·211 cm.²/sec. at 15° C. and 76 cm. pressure.

Using this value of D_{11}, the value of the constant k in the self-diffusion relation $D_{11} = \frac{k\eta}{\rho}$ may now be calculated:

$$\eta = 1{\cdot}737 \times 10^{-4} \text{ c.g.s. units at } 15^\circ \text{ C. (Smith 1922),}$$

$$\rho = 1{\cdot}184 \times 10^{-3} \text{ g./c.c. at } 15^\circ \text{ C. and 76 cm.}$$

Substituting these values gives $k = 1{\cdot}44$.

The mean value of D_{12} for nitrous oxide and carbon dioxide (D_{11} for either gas) was found to be 0·107 cm.²/sec. at 15° C. and 76 cm. pressure. Also

$$\eta = 1{\cdot}441 \times 10^{-4} \text{ c.g.s. units at } 15^\circ \text{ C. (Smith 1922),}$$

$$\rho = 1{\cdot}874 \times 10^{-3} \text{ g./c.c. at } 15^\circ \text{ C. and 76 cm.}$$

Substituting these values in the self-diffusion relation gives $k = 1{\cdot}39$.

Thus in both the cases of self-diffusion investigated experimentally the value of the constant k lies between the values given by the classical theory for Maxwellian molecules and rigid elastic spheres. The value of k for nitrous oxide and carbon dioxide is smaller than the value for nitrogen and carbon monoxide.

The calculation of the coefficients of self-diffusion of nitrogen and carbon dioxide from the results given in Tables V, VI and VII is carried out as follows (Kelvin 1904). Maxwell's formula based on the rigid elastic sphere model is

$$D_{12} = \frac{1}{2\sqrt{(6\pi)}} \frac{V}{N} \frac{1}{S_{12}^2} \sqrt{\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right)},$$

where ω_1 and ω_2 are the masses of the molecules referred to hydrogen as unity, $S_{12} = \frac{S_1 + S_2}{2}$ = the mean of the radii of the two kinds of molecules, V = the root mean square velocity of a hydrogen molecule at 0° C., N = the number of molecules in a cubic centimetre of gas at 0° C. and 76 cm. pressure.

Let
$$\frac{1}{2\sqrt{(6\pi)}} \frac{V}{N} = \text{a constant} = K^2;$$

then
$$D_{12} = \frac{K^2}{S_{12}^2} \sqrt{\left(\frac{1}{\omega_1} + \frac{1}{\omega_2}\right)}.$$

Now if D_{12} and ω_1 and ω_2 are known K^2/S_{12}^2 can be calculated. If this is done for the three pairs of gases then S_1/K may be determined, for

$$\frac{S_1}{K}=\frac{S_{12}}{K}+\frac{S_{13}}{K}-\frac{S_{23}}{K}$$

whence by substitution $$D_{11}=\frac{K^2}{S_1^2}\sqrt{\frac{2}{\omega_1}}.$$

The results of this calculation are given in Tables VIII, IX and X.

TABLE VIII

Gas	Mol. wt. O = 16·000	ω $H_2 = 1$	$\sqrt{\frac{2}{\omega}}$
(1) H_2	2·015	1·000	1·414
(2) N_2	28·016	13·90	0·379
(3) CO_2	44·000	21·83	0·303

TABLE IX

Gas mixture	D_{12} exp.	$\sqrt{\left(\frac{1}{\omega_1}+\frac{1}{\omega_2}\right)}$	S_{12}/K
(1)-(2) H_2—N_2	0·743	1·0354	1·180
(2)-(3) N_2—CO_2	0·158	0·3431	1·472
(3)-(1) CO_2—H_2	0·619	1·0227	1·286

TABLE X

Gas	S_1/K	D_{11}
H_2	0·994	1·43
N_2	1·366	0·203
CO_2	1·578	0·121

It will be seen that there is reasonable agreement between the values D_{11} obtained directly for the pairs of gases nitrogen-carbon monoxide and nitrous oxide-carbon dioxide, and for D_{11} obtained indirectly by the Kelvin method for nitrogen and for carbon dioxide. The diffusing molecules for the pairs mentioned are in each case similar in many respects, and the result obtained is in accordance with the predictions of the classical kinetic theory. As the molecules must be distinguishable for a diffusion measurement to be made it is not possible to devise an experiment to examine the quantum theory conception of self-diffusion. The results are therefore not opposed to the conclusions of Massey and Mohr, and in fact were to be expected from their work.

We thank Professor S. W. J. Smith for providing full facilities for this work in the Physics Department of the University of Birmingham. We are also very much indebted to Dr. T. L. Ibbs for bringing this problem (which was originally suggested by Professor S. Chapman) to our notice; and for much help and advice in carrying out the work.

SUMMARY

Diffusion experiments were carried out with the two pairs of gases, nitrogen-carbon monoxide and nitrous oxide-carbon dioxide. Previous experiments have shown that each pair consists of two gases whose physical properties are almost exactly the same. The diffusion experiments made with these pairs of gases are therefore, from the point of view of the classical kinetic theory, a close approximation to self-diffusion. The coefficients of self-diffusion for nitrogen and for carbon dioxide were also calculated from experiments on the pairs of gases: hydrogen-carbon dioxide, hydrogen-nitrogen, and nitrogen-carbon dioxide. There is good agreement between the coefficients obtained by the two methods. The result is not contrary to the conclusions reached in the quantum theory of collisions of Massey and Mohr, as from that point of view it is necessary for the molecules in true self-diffusion to be identical and indistinguishable.

REFERENCES

Chapman, S. 1916 *Philos. Trans.* A, **216**, 279–348.
— 1918 *Philos. Trans.* A, **217**, 115–98.
Daynes, H. A. 1920 *Proc. Roy. Soc.* A, **97**, 273–86.
Handbuch der Physik, 2nd ed. **1**, 1415.
Harteck and Schmidt 1933 *Z. phys. Chem.* B, **21–22**, 447–58.
Ibbs, T. L. and Underwood, L. 1927 *Proc. Phys. Soc.* **39**, 227–37.
Jeans, J. H. 1921 "Dynamical Theory of Gases", 3rd ed. p. 284. Camb. Univ. Press.
Kelvin 1904 "Baltimore Lectures", p. 293. London: C. J. Clay and Sons.
Lonius, A. 1909 *Ann. Phys., Lpz.*, **29**, 664–78.
Loschmidt 1870 *S.B. Akad. Wiss. Wien.*, **61**, 367; **62**, 468.
Massey, H. S. W. and Mohr, C. B. O. 1933 *Proc. Roy. Soc.* A, **141**, 434–53.
Obermayer 1880 *S.B. Akad. Wiss. Wien*, **81**, 1102.
Schmidt, R. 1904 *Ann. Phys., Lpz.*, **14**, 801–21.
Smith, C. J. 1922 *Proc. Phys. Soc.* **34**, 155–65.

Neutrons, Degeneracy and White Dwarfs

BY D. S. KOTHARI, PH.D. (Cambridge)
Department of Physics, University of Delhi

(*Communicated by M. N. Saha, F.R.S.—Received* 8 *June* 1937)

SUMMARY AND INTRODUCTION

The present paper gives an account of work which, as a consequence of the fact that the neutron mass is greater than the combined mass of the proton and the electron, leads to (i) the existence of an *upper limit* for the pressure exerted by *degenerate* electron gas (present in cold† ionized matter), and (ii) certain interesting conclusions regarding the presence of hydrogen in the interior of a white dwarf star and its relation to the radius of the star. It will be proved (in a straightforward way) that a white dwarf having a mass greater than a certain critical mass M_0 (about $4\odot/\mu^2$, where $\odot$ is the mass of the sun) cannot contain hydrogen in its interior. In fact, it seems possible, as will appear in the sequel, to determine the neutron-proton mass difference from purely astrophysical data regarding the white dwarfs.‡

1—The recent work§ of Rutherford and others on nuclear disintegration and of Aston on mass-spectrograph measurements has now established the fact that the neutron mass exceeds the combined mass of proton and electron by 0·0009 mass unit, and it is improbable that the error can exceed 0·0001 unit. We shall denote this neutron mass excess by U, and in numerical work take

$$U = 0{\cdot}001 \text{ mass unit} = 1{\cdot}48 \times 10^{-6} \text{ erg} = 0{\cdot}93 \times 10^{6} \text{ e-volt}.$$

---

† The word *cold* is used in a technical sense. Matter will be referred to as *cold* or *degenerate* if its temperature and density are such that any free electrons present constitute a degenerate gas in the sense of Fermi-Dirac statistics. The stellar material in the interior of white dwarfs and planets will, therefore, be referred to as *cold*. The existence of an upper limit for the radius of a *cold* body follows from the theory of pressure ionization (see Kothari and Majumdar 1936*a*; Kothari 1936).

‡ The main results contained in this paper were submitted in an abstract form to the Indian Science Congress (in October 1936) and printed in their Abstracts (**1**, p. 1, 1937). While the present paper was being written for the Press, the writer came across a most interesting article by Hund (1936) on "Materie unter sehr hohen Drucken und Temperaturen", where essentially the same result as (i) above has been given, and he desires to acknowledge in a general way his indebtedness to this article by Hund. The astrophysical applications contained in the present paper are, I believe, new.

§ A short account is given in Feather (1936).

The existence of an upper limit for the pressure exerted by degenerate electron gas is a direct consequence of the above fact, as can be seen in the following way.

Let us consider degenerate matter, e.g. a mixture of free electrons and protons, and suppose this to be gradually compressed (density increased) so that E^*, the maximum energy of the Fermi distribution (which in the *non-relativistic* case is 5/3 times the mean residual or null-point energy† per electron) attains the value U. Let the corresponding value for the pressure of the degenerate electron gas be denoted by p_m. Any further compression (i.e. increase in density) will tend to increase E^* above U and endow some of the electrons with kinetic energy greater than U, but these electrons, *as they have an energy greater than* U, will combine with protons transforming them into neutrons, the neutrinos being furnished from negative energy states according to the process pictured by Bloch and Møller (1935), also Bethe and Bacher (1936).‡ Therefore, electrons with kinetic energy greater than U do not exist in a free state, but are removed by the protons to form neutrons with the result that E^* will not exceed U, and consequently the pressure of the degenerate electron gas will not exceed p_m.

2—Our discussion so far has been in general terms. We shall now substitute the proper numerical value for U and evaluate p_m.

The relation connecting pressure and electron concentration for a degenerate gas taking account of relativity has been worked out by several investigators (Majumdar 1932). The pressure is given by

$$p = \frac{8\pi m^4 c^5}{3h^3}\{\tfrac{1}{8}x(1+x^2)^{\frac{1}{2}}(2x^2-3)+\tfrac{3}{8}\log[x+(1+x^2)^{\frac{1}{2}}]\}, \qquad (1)$$

and E^* is given by $$E^* = mc^2[(1+x^2)^{\frac{1}{2}}-1], \qquad (2)$$

---

† Expressions for the null-point energy, null-point pressure, and other allied quantities are given in a tabular form by Kothari (1931).

‡ It may be noted, as has been pointed out by these investigators, that the probability of an electron (which has energy in excess of U) transforming a proton into a neutron is so small that the process can scarcely be detected under the usual laboratory conditions. However, this does not diminish its importance from the point of view of stellar applications. It will be seen (by putting the order of magnitudes of the various quantities in the formulae given by the above workers) that in the interior of a white dwarf having density large enough for the free electrons to have an energy greater than U, a proton will remain free, i.e. not suffer transformation into a neutron, for a time negligibly small compared with the "life" of the star, and it is really this ratio of the "life of free proton" to the "life of star" that matters in stellar applications of the theory.

where x stands for
$$x = \frac{h}{mc}\left(\frac{3n}{8\pi}\right)^{\frac{1}{3}}, \tag{3}$$
m denotes the electron rest-mass, n the number of electrons per unit volume, and c and h have their usual meanings.

If we put $E^* = U = 1{\cdot}48 \times 10^{-6}$ erg, we obtain from (2)
$$x = 2{\cdot}65. \tag{4}$$
Substituting this value of x in (3) we get for n_m the maximum possible free electron concentration the value
$$n_m = 1{\cdot}09 \times 10^{31} \text{ per cm.}^3, \tag{5}$$
and from (1) we have
$$p_m = 5{\cdot}20 \times 10^{24} \text{ dynes per cm.}^2. \tag{6}$$

We notice from equation (2) that for $x \ll 1$, $E^* \ll mc^2$, i.e. the maximum kinetic energy is less than the rest-mass energy, and thus for this case of $x \ll 1$, the relativity correction is negligible. It is called the *non-relativistic* case and (1) and (2) approximate to
$$p = \frac{8\pi}{15}\frac{h^2}{m}\left(\frac{3n}{8\pi}\right)^{\frac{5}{3}}, \tag{7}$$
$$E^* = \frac{h^2}{2m}\left(\frac{3n}{8\pi}\right)^{\frac{2}{3}}. \tag{8}$$
Expressing p in terms of E^* we have
$$p = \frac{32\sqrt{2}\pi}{15}\frac{m^{\frac{3}{2}}}{h^3}E^{*\frac{5}{2}}. \tag{9}$$
When $x \gg 1$, $E^* \gg mc^2$: it is called the *relativistic case*. For this case (1) and (2) reduce to
$$p = \frac{2\pi}{3}hc\left(\frac{3n}{8\pi}\right)^{\frac{4}{3}}, \tag{10}$$
$$E^* = hc\left(\frac{3n}{8\pi}\right)^{\frac{1}{3}}, \tag{11}$$
and corresponding to (9) we have
$$p = \frac{2\pi}{3}\frac{E^{*4}}{h^3c^3}. \tag{12}$$

If we substitute $E^* = U = 1{\cdot}48 \times 10^{-6}$ erg in (9) and (12) we obtain for p_m the values $2{\cdot}42 \times 10^{24}$ and $1{\cdot}34 \times 10^{24}$ dynes/cm.^{2} respectively, which may be compared with the value $5{\cdot}20 \times 10^{24}$ obtained from the exact equation (1).

3—In the foregoing calculations we have considered the *cold* matter to be composed of protons and electrons. We have now to consider the modification in the value of p_m when the degenerate matter is composed of free electrons and nuclei other than protons. Consider, for example, a mixture of free electrons and deuterons. The mass of the deuterium atom is 2·0147, and thus for a free electron to be able to transform a deuteron into two neutrons, the electron must have kinetic energy equal to the mass difference between two neutrons and a deuteron, i.e. 0·0035 mass unit. In the case of other elements, a reference to the table of atomic masses shows that for an electron to transform a nucleus of mass number M and atomic number Z into a neutron and another nucleus of mass number $(M-1)$ and atomic number $(Z-1)$, it must have kinetic energy round about the value 0·01 mass unit.

Thus we see that in a mixture of degenerate electron gas and nuclei (M, Z), a free electron cannot have a kinetic energy greater than 0·01 mass unit, for if it has an energy greater than this amount, it will combine with the nucleus and transform it into a neutron and another nucleus $(M-1, Z-1)$. In order, therefore, to estimate the maximum values for the electron concentration and pressure for this case, we have to substitute $E^* = 0{\cdot}01$ mass unit in (2), and the corresponding value of x in (1). As, however, $E^* = 0{\cdot}01$ is about 20 times the electron rest mass ($= 0{\cdot}000547$ mass unit), it will be sufficiently accurate to use the *relativistic* formula (10) and (11) instead of the exact equations (1) and (2). Equation (11) gives for n'_m the maximum electron concentration

$$n'_m = 3{\cdot}61 \times 10^{33} \text{ per cm.}^3,$$

and equation (10) gives for p'_m the maximum electron pressure

$$p'_m = 1{\cdot}34 \times 10^{28} \text{ dynes/cm.}^2.$$

4—Before we consider the application of the above results to the white dwarf stars, we shall summarize in this section the essential theoretical results concerning their structure which are relevant to the present discussion. In the case of the white dwarf stars the researches beginning with Fowler and later developed extensively by Milne and others have shown that the essential features of their constitution can be explained by considering the stellar material in their interiors to be degenerate. Further, as has been proved in a recent paper (Kothari 1936) *in the case of the white*

dwarf stars the degenerate matter will be fully ionized, i.e. the atoms reduced to free electrons and bare nuclei. A white dwarf is characterized by a low luminosity, a high effective temperature and large mean density. In preliminary theoretical treatment it is usual to assume the luminosity of a white dwarf to be zero, but as has been shown elsewhere (Kothari 1932), the small luminosity of a white dwarf has little effect on the calculated values of its radius and mean density.

The radius R of a star composed of *cold* matter is connected with its mass M by the relation† (neglecting relativity effects, this being justified so long as the mass is not much larger than the sun)

$$R = \frac{5(\omega_{\frac{3}{2}}^{\circ})^{\frac{1}{3}}}{2^{\frac{5}{3}}\pi^{\frac{1}{3}}} \cdot \frac{K}{G\mu^{\frac{5}{3}}} \frac{1}{M^{\frac{1}{3}}}$$

$$= \frac{2{\cdot}79 \times 10^{9}}{\mu^{\frac{5}{3}}} \left(\frac{\odot}{M}\right)^{\frac{1}{3}} \text{cm.,} \tag{13}$$

where $\odot$ is the mass of the sun, μ is the mean molecular weight per free electron and K is the "degenerate gas constant".

As μ is defined to be the average weight (on the *chemical scale*) per electron, $\rho/\mu m_{\mathrm{H}}$—where ρ denotes the density and m_{H} is the mass of the hydrogen atom—gives the number of free electrons per unit volume. If ρ_{mean} denotes the mean density, then from (13) we have

$$\rho_{\text{mean}} = \frac{M}{\frac{4\pi}{3} R^3} = \frac{96\pi}{125(\omega_{\frac{3}{2}}^{\circ})} \left(\frac{G}{K}\right)^3 \mu^5 M^2$$

$$= 2{\cdot}19 \times 10^4 \mu^5 \left(\frac{M}{\odot}\right)^2 \text{g./cm.}^3, \tag{14}$$

and this is connected with the central density ρ_c by the relation (given in Milne's notation)

$$\frac{\rho_c}{\rho_{\text{mean}}} = \frac{(\sigma_{\frac{3}{2}})^{\frac{3}{2}}}{3(\omega_{\frac{3}{2}}^{\circ})} = 5{\cdot}99; \tag{15}$$

---

† Milne (1932). The notation used here is slightly different from Milne's paper. Our μ denotes the mean molecular weight (per electron) on the *chemical scale*, whereas Milne's μ (mean molecular weight per particle, i.e. electron as well as the "heavy particle") is in grams. Again our $K/\mu^{\frac{5}{3}}$ is his K. This is done to show explicitly the dependence on μ. The "degenerate gas constant" K in the non-relativistic case is defined by the relation: pressure of degenerate gas $= K\rho^{\frac{5}{3}}/\mu^{\frac{5}{3}}$, where ρ is the density of the degenerate matter.

and thus we have

$$\rho_c = \mu m_{\mathrm{H}} n_c = \frac{32\pi(\sigma_{\frac{3}{2}})^{\frac{1}{2}}}{125(\omega^{\circ}_{\frac{3}{2}})^2}\left(\frac{G}{K}\right)^3 \mu^5 M^2$$

$$= 1{\cdot}31 \times 10^5 \mu^5 \left(\frac{M}{\odot}\right)^2 \text{g./cm.}^3, \tag{16}$$

or

$$\frac{M}{\odot} = 2{\cdot}76 \times 10^{-3}\frac{(m_{\mathrm{H}} n_c)^{\frac{1}{2}}}{\mu^2}. \tag{17}$$

On eliminating M between (13) and (16) we obtain

$$R = \frac{5^{\frac{1}{2}}}{2^{\frac{3}{2}}\pi^{\frac{1}{2}}}(\sigma_{\frac{3}{2}})^{\frac{1}{2}}\left(\frac{K}{G}\right)^{\frac{1}{2}}\frac{1}{(m_{\mathrm{H}} n_c)^{\frac{1}{6}}}\frac{1}{\mu}. \tag{18}$$

5—We are now in a position to consider the application of the results of § 2. Equation (16) shows that n_c, the free-electron concentration at the centre of a white dwarf star, depends directly on the square of the mass of the star. Let M_0 denote that mass for which the central electron concentration has the value $n_m(= 1{\cdot}09 \times 10^{31})$ *which represents the maximum value of electron concentration possible in "cold" matter which contains protons*. For an electron concentration greater than n_m implies the existence of free electrons possessing kinetic energy greater than 0·001 mass unit, but such electrons cannot exist in a free state in the presence of protons—they will combine with protons to form neutrons.

Substituting $n_c = n_m$ in (17), we obtain for M_0

$$M_0 = \frac{11{\cdot}7}{\mu^2}\odot, \tag{19}$$

and the radius for this mass is given by (18),

$$R_0 = \frac{1{\cdot}23 \times 10^9}{\mu}\text{ cm.} \tag{20}$$

Equations (19) and (20) are based on (13) which, as already mentioned, does not take account of the relativity effect. The relativistic correction is unimportant for masses which are less than or comparable with the solar mass, but as in our case M_0 is appreciably larger than $\odot$, the relativistic correction will also be appreciable. A reference to Table III in Chandrasekhar's paper (Chandrasekhar 1935), where the problem of the equilibrium configurations for mass M when the equation of state of the matter composing it is defined by (1) has been investigated in complete detail, shows that the *relativistic* effect will lower the value of M_0 by a factor of about 3, i.e. M_0 becomes about $\frac{4\odot}{\mu^2}$, but the change in R_0 is very little.

The value M_0 represents a critical mass in the following sense. If the stellar mass is smaller than M_0, the central electron concentration n_c which varies as the square of the stellar mass, will be less than n_m, but for $M > M_0$, n_c will exceed n_m. This shows that for $M < M_0$, free electrons and protons can exist together without leading to neutron formation, but for $M > M_0$ (for the stellar configuration to be in equilibrium under its own gravitational forces) the central electron concentration must exceed n_m and therefore free protons cannot be present. The essential point is this, that in a degenerate gas the maximum electron energy depends on electron concentration, and so long as protons are present, this maximum kinetic energy cannot exceed 0·001 mass unit—for electrons with energy greater than this will unite with protons and transform them into neutrons. This in turn implies that the electron concentration must be less than $n_m = 1{\cdot}09 \times 10^{31}$ per cm.^{3}, and this is only possible for a stellar mass less than (or equal to) M_0.

6—There remains now the important question: What will be the observational consequences of the theoretical result that white dwarfs can be classified as (1) hydrogen-containing, $M < M_0$ and (2) non-hydrogen containing, $M > M_0$?

It seems that there are three ways in which the effect of the hydrogen content of the interior will manifest itself.

(i) If conditions be such that no hydrogen can exist in the interior of a star, then it is probable that there will be no appreciable amount of it in the outer portions also, for if any hydrogen be present in the outer portions it will diffuse towards the interior (which contains no hydrogen) where it will be "destroyed" and the process will continue till the outer layers have lost almost all their hydrogen.

(ii) If no hydrogen is present in the interior then it is reasonable to postulate that the rate of energy generation by nuclear transformation will be very different from what it would be in hydrogen-containing white dwarfs. As is now well known most of the subatomic energy is released during the transformation of hydrogen into heavier elements, and an absence of hydrogen would imply a very small (or negligible) rate of energy generation compared to the hydrogen-containing white dwarfs. Possibly the non-hydrogen-containing white dwarfs would be too faint to be observationally accessible.

(iii) The value of the average electron molecular weight μ will be different for the two types of stars. For the non-hydrogen containing star μ will exceed 2, but for the hydrogen-containing star the value will depend largely on the proportion of hydrogen present—for a fully ionized mixture of 30 % by weight hydrogen and 70 % iron μ is about 1·6.

Coming now to the known white dwarfs, all of them show hydrogen lines and roughly fall in the same class from the point of view of their L/M values. If we *assume* them all to fall in the hydrogen-containing class, then our theory predicts that their radii must be less than

$$R_0 = \frac{1{\cdot}23 \times 10^9}{\mu} \text{ cm.},$$

or, putting $\mu = 2$, which is possibly its maximum value, R_0 equals about the radius of the earth ($6{\cdot}4 \times 10^8$ cm.). This conclusion is somewhat at variance with Kuiper's recent discovery (Kuiper 1935) of a white dwarf A.C. (+7·0°, 8247). Kuiper estimates its radius to be between $(4{\cdot}3 - 2{\cdot}7) \times 10^8$ cm., and it is considered to be the smallest white dwarf so far discovered. At the present stage of our theory it is hardly justifiable to discuss this—certainly not large—disagreement at length.

It is only necessary in conclusion to point out that the result, that white dwarfs above a certain critical mass cannot contain hydrogen in their interiors, follows from the theory in quite a straightforward way; the speculative element enters when we discuss how this absence of hydrogen in the interior will affect the observational characteristics of the white dwarfs. And when this later part is put on a somewhat more secure basis, we shall have a purely astrophysical method of determining the neutron mass.

REFERENCES

Bethe and Bacher 1936 *Rev. Mod. Phys.* **8**, 82–229. (See p. 196.)
Bloch and Møller 1935 *Nature, Lond.*, **136**, 987.
Chandrasekhar 1935 *Mon. Not. R. Astr. Soc.* **95**, 207–25.
Feather 1936 "Nuclear Physics", Chap. v. Camb. Univ. Press.
Hund 1936 *Ergebn. Exakt. Naturwiss.* **15**.
Kothari 1931 *Phil. Mag.* **12**, 665–76.
— 1932 *Mon. Not. R. Astr. Soc.* **93**, 61–90.
— 1936 *Mon. Not. R. Astr. Soc.* **96**, 833–43.
Kothari and Majumdar 1936*a* *Nature, Lond.*, **137**, 157–8.
— — 1936*b* *Proc. Nat. Acad. Sci. U. P. India*, **6**, 57–65.
Kuiper 1935 *Pub. Astr. Soc. Pacific*, **47**, 307–10.
Majumdar 1932 *Astr. Nachr.* **247**, 248.
Milne 1932 *Mon. Not. R. Astr. Soc.* **92**, 611–43.

Nuclear Excitation and Disintegration Collisions involving Strong Interaction—I

By H. S. W. Massey, Ph.D., *Independent Lecturer in Math. Physics, Queen's University, Belfast*
and C. B. O. Mohr, Ph.D., *Lecturer in Physics, University of Cape Town*

(*Communicated by R. H. Fowler, F.R.S.—Received* 25 *June* 1937)

The study of collisions between nuclear particles has developed to a remarkable extent with the discovery of the neutron and the introduction of artificial methods for effecting nuclear disintegration. It has been found in the last few years that the interpretation of the observed results is by no means as simple as was first expected. This situation is most apparent when the explanation of the variation of probability of capture of slow neutrons by different nuclei is considered. This probability varies in a very irregular manner from element to element and pronounced selective effects occur in certain cases. Attempts to explain (Elsasser and Perrin 1935; Bethe 1935) these resonance phenomena in terms of the usual approximations of quantum collision theory were soon found to be inadequate. All such attempts were based on the assumption that the chance of a nuclear collision being elastic is high compared with that of its resulting in capture or excitation. A high probability of capture (with emission of radiation) or excitation could then only appear together with a high probability of elastic collision and this is frequently contradicted by the experimental results. The sharpness of the observed resonance phenomena was also difficult to understand on this basis.

It was first pointed out by Bohr (1936) that the initial assumptions concerning the probability of elastic collisions, virtually involving the treatment of the elastic scattering as a one-body problem in the first approximation, cannot be valid for nuclei in which the particles, even if existing separately in the nuclei at all, are so closely packed. On making a close collision with a nucleus a particle, such as an α-particle, neutron or proton, comes into close and strong interaction with a number of nuclear particles and its incident energy becomes distributed among them. It is only when a particular particle receives sufficient energy to leave the quasi-stable complex formed that a disintegration particle is emitted. (This may of course be the original incident particle, in which case the collision would be an elastic or excitation one.) Otherwise the surplus energy is emitted as radiation. The resonance phenomena arise from the energy levels of the quasi-stable complex. If the incident energy is such that the total energy is equal or nearly equal to that of one of these energy levels, the range of interaction and hence the collision cross-section is quite large. This point of view must be adopted not only when dealing with neutron collisions but in all cases in which the impinging particle does not possess an energy greatly in excess of the minimum necessary for the process to occur. Disintegrations produced by charged particles, in which resonance effects have been observed for some time (Feather 1937, p. 154), must therefore be capable of description in this way.

Although there is little doubt of the correctness of Bohr's explanation it is not easy to place it on anything like a quantitative basis. Breit and Wigner (1936) have discussed certain quantitative aspects of the theory concerning the emission of radiation employing the usual methods of radiation theory, and Bethe and Placzek (1937), using much the same methods, have dealt in general terms with the whole range of phenomena to be expected, but there has been no attempt made to deal with the problem in terms of quantum collision theory worked out more accurately than attempted hitherto (not assuming the prior importance of elastic scattering). It is very unlikely that such a mode of approach could ever be employed to deal with collisions between nuclei involving a great number of particles and it is ill-adapted for the discussion of radiative collisions in any case, but it is by no means impossible to consider in this way collisions involving relatively few particles (such as those between protons, neutrons or α-particles and the nuclei in the first row of the periodic table) in which no radiation is emitted. These represent the interesting transition cases where the more usual methods have not completely lost their validity and it should be possible to derive specific knowledge of the properties of light nuclei by refining the usual approximate methods. As the mathematical and computational difficulties which must be faced are considerable the first step to take is to deal as accurately as possible with problems of a three-body nature of the type

$$\alpha\text{-particle} + (\text{proton} + \text{nuclear ``core''}) \rightarrow \text{proton} + (\alpha\text{-particle} + \text{nuclear core})$$

in which the nuclear "core" is treated as a single particle. In this paper we describe in detail the results of such calculations for excitation and disintegration phenomena.

The essential feature of Bohr's theory is the impossibility of obtaining an accurate approximation by regarding the elastic scattering as most important and so we have worked out in detail the exact solutions for certain cases without assuming predominance of elastic collisions. To obtain such solutions the specifications of the colliding system have been somewhat idealized but without loss of the essential features. Numerical solutions have been obtained for two cases of excitation, for disintegration by an α-particle with neutron emission and for disintegration by a slow neutron involving emission of a charged particle.

From these solutions it is possible to derive information concerning the way in which the usual approximation, which considers the resonance effects to be due to the energy of the incident particle coinciding with a "virtual" level in the average potential field acting on the particle, needs modification. Although the methods employed are capable of extension to three- and higher-body problems, such extension will involve greatly increased labour and the results already obtained for the two-body cases are of sufficient interest to warrant description before the more complicated cases have been worked out. It is of interest to notice also that the problems concerned in this paper are also of importance in the discussion of other collision problems involving strong interaction.

1—GENERAL THEORETICAL CONSIDERATIONS

Let us consider a collision between two nuclei. The motion of the centre of gravity can be separated out so that the collision can be described by a function Ψ of the co-ordinates determining the relative position of the particles comprising the nuclei. We shall consider for simplicity the case where the impinging nucleus is a single particle; the more general case presents no new features.

The choice of defining co-ordinates is dictated largely by the nature of the particular collision in which we are interested. The discussion of an ordinary inelastic collision in which no rearrangement of particles between the two systems occurs follows on somewhat different lines to that of a collision involving capture, so we give them separate consideration. The methods employed are nevertheless similar in essentials.

(a)—Excitation Collisions

Here we choose as co-ordinates the relative co-ordinates $\mathbf{r}$ of the colliding nucleus and the centre of mass of the struck nucleus, and those, $\mathbf{s}$, of the particles comprising this nucleus also referred to its centre of mass. The wave equation in these co-ordinates takes the form

$$\left[-\frac{h^2}{8\pi^2 M_1}\nabla^2 + H(\mathbf{s}) - E + V(\mathbf{r}, \mathbf{s})\right]\Psi = 0, \tag{1}$$

where M_1 is the reduced mass, H the Hamiltonian operator for the motion of the bound nuclear particles relative to their centre of mass and $V(\mathbf{r}, \mathbf{s})$ the interaction energy. Ψ, being a proper function of the co-ordinates $\mathbf{s}$, can be expanded in a series of the orthogonal set of proper functions ψ which satisfy

$$(H - E_n)\psi_n = 0, \quad n = 0, 1, 2, \dots. \tag{2}$$

Writing then

$$\Psi = \Sigma_n F_n(\mathbf{r})\,\psi_n(\mathbf{s}), \tag{3}$$

we find

$$\left[\frac{h^2}{8\pi^2 M_1}\nabla^2 + (E - E_n)\right]F_n = \int \psi_n^*(\mathbf{s})\,V(\mathbf{r}, \mathbf{s})\,\Psi(\mathbf{r}, \mathbf{s})\,d\mathbf{s}, \quad n = 0, 1, 2\dots, \tag{4}$$

an exact set of equations. To describe the collision solutions of these equations must be obtained which are proper functions and have asymptotic forms

$$\begin{aligned} F_0 &\sim e^{ikz} + f(\theta, \phi)\,e^{ikr}r^{-1}, \\ F_n &\sim f_n(\theta, \phi)\,e^{ik_n r}r^{-1}, \quad n = 1, 2, \dots; \end{aligned} \tag{5}$$

where

$$\frac{h^2}{8\pi^2 M_1}k_n^2 = E - E_n, \quad \frac{h^2}{8\pi^2 M_2}k^2 = E - E_0, \tag{6}$$

and it is supposed that the struck nucleus is initially in its normal state. The cross-section for an elastic collision is then

$$Q_0 = \iint |f(\theta, \phi)|^2 \sin\theta\,d\theta\,d\phi, \tag{7}$$

and for an inelastic collision in which the nth state of the struck nucleus is excited

$$Q_n = \iint |f_n(\theta, \phi)|^2 \sin\theta\,d\theta\,d\phi. \tag{8}$$

Before examining approximate methods of solving these equations it is convenient to consider the form (3) for Ψ substituted on the right-hand side of the equations (4) giving

$$(\nabla_2+k_n^2)F_n = \frac{8\pi^2M_1}{h^2}\Sigma_t V_{nt}F_t, \quad n=0,1,2,\ldots, \tag{9}$$

where

$$V_{nt} = \int \psi_n^*(\mathbf{s})\, V(\mathbf{r},\mathbf{s})\, \psi_t(\mathbf{s})\, d\mathbf{s}. \tag{10}$$

Neglecting all non-diagonal matrix elements V_{nt}, $n \neq t$, except those connected with the incident waves gives the set of equations

$$\left(\nabla^2+k^2-\frac{8\pi^2M_1}{h^2}V_{00}\right)F_0 = 0, \tag{11}$$

$$\left(\nabla^2+k_n^2-\frac{8\pi^2M_1}{h^2}V_{nn}\right)F_n = V_{n0}F_n, \quad n\neq 0, \tag{12}$$

which may be solved by the method of Green's functions in terms of the solutions of the equations with right-hand side zero. This is the "method of distorted waves" (Mott and Massey 1933, p. 100) and requires for its validity that the inelastic scattering be small compared with the elastic (unless both are so small that F_0 may be represented effectively by a plane wave). The matrix element V_{n0} must be so small that $V_{n0}F_n$ is small compared with the terms on the left-hand side of (11). We see also that the solution obtained by this method does not automatically satisfy the conservation of charge and therefore gives results which are very far from the truth when V_{n0} is large.

Further, in nuclei containing a large number of particles, the number of quantum states is very great and the matrix elements V_{nt}, even if small, give a far from unimportant contribution to the right-hand side of (9). We do not propose to deal with this type of failure of the distorted wave approximation in this paper, but instead to examine the behaviour to be expected when the method fails owing to F_0 or F_n being largely concentrated within nuclear dimensions even when the number of important excited states is limited. We therefore take the simplest case of one excited state only and so have to deal with the simultaneous equations:

$$\nabla^2F_0+\left(k^2-\frac{8\pi^2M_1}{h^2}V_{00}\right)F_0 = V_{01}F_1, \tag{13}$$

$$\nabla^2F_1+\left(k_1^2-\frac{8\pi^2M_1}{h^2}V_{00}\right)F_1 = V_{10}F_0, \tag{14}$$

without assuming V_{01} small. These equations necessarily give solutions satisfying the conservation of charge, and these may be taken as representative of the true solutions for more complicated cases in a way which could not be expected from the method of distorted waves. Description of the detailed calculations and results will be postponed until we have considered capture collisions as far as possible on the same footing.

(*b*)—*Collisions involving Capture*

The possibility of capture of the incident particle by the nucleus is represented in the preceding theory by the existence of solutions F_n for which k_n^2 is negative. If the corresponding wave function ψ_n is one of the continuous spectrum, the capture is accompanied by particle disintegration. It is difficult to obtain any quantitative results concerning capture cross-sections in this way, but it does offer some prospect of dealing with the effect of the possibility of capture on elastic scattering even when the calculation of the capture cross-sections themselves may be intractable. However, we defer such considerations to a later paper and proceed to discuss capture with particle emission from a modified view point.

We restrict ourselves to cases where the incident particle is absorbed and a second particle emitted. The co-ordinates which are most suitable are now the relative co-ordinates $\boldsymbol{\rho}$ of the centres of mass of the two nuclei and those $\boldsymbol{\sigma}$ of the nuclear particles which are bound after the collision, referred to their common centre of mass. The wave function Ψ of (*a*) can then be regarded as a function of $\boldsymbol{\rho}$, $\boldsymbol{\sigma}$ and a solution of the equation

$$\left[-\frac{h^2}{8\pi^2 M_2}\nabla_\rho^2 + H'(\sigma) - E + U(\boldsymbol{\rho}, \boldsymbol{\sigma})\right]\Psi = 0, \tag{15}$$

where H' is the Hamiltonian for the internal motion in the final composite nucleus, U the interaction energy between the final nuclei and M_2 the final reduced mass. Further, Ψ may be expanded in terms of the orthogonal set of wave functions ϕ_ν satisfying

$$(H' - E_\nu)\,\phi_\nu = 0. \tag{16}$$

Hence, writing

$$\Psi = \Sigma_\nu G_\nu(\boldsymbol{\rho})\,\phi_\nu(\boldsymbol{\sigma}), \tag{17}$$

we obtain as before

$$\left[\frac{h^2}{8\pi^2 M_2}\nabla_\rho^2 + E - E_\nu\right] G_\nu = \int \phi_\nu^*(\boldsymbol{\sigma})\, U(\boldsymbol{\rho}, \boldsymbol{\sigma})\,\Psi\, d\boldsymbol{\sigma}, \quad \nu = 0, 1, 2, \ldots. \tag{18}$$

If the solution is to describe the collision, we must have

$$G_\nu \sim e^{-1} e^{i\kappa_\nu \rho} g_\nu(\vartheta, \chi), \tag{19}$$

where

$$\kappa_\nu^2 = \frac{8\pi^2 M_2}{h^2}(E - E_\nu),$$

while for large r, Ψ must have the form

$$\Psi \sim e^{ikz}\psi_0(\mathbf{s}) + \text{terms } O(r^{-1}), \tag{20}$$

corresponding to outgoing spherical waves. The cross-section for the particular capture process which leaves the composite nucleus in state ν is then

$$Q_c^\nu = \frac{\kappa_\nu}{k}\iint |g_\nu(\vartheta, \chi)|^2 \sin\vartheta\, d\vartheta\, d\chi. \tag{21}$$

It is clearly impossible to solve the equations (18) exactly, and in practice it is usual to employ a method of approximation similar to that of "distorted waves" sketched out in (*a*), giving

$$\left[\nabla_\rho^2 + \kappa_\nu^2 - \frac{8\pi^2 M_2}{h^2} U_{\nu\nu}\right] G_\nu = \frac{8\pi^2 M_2}{h^2}\int \phi_\nu^*(\boldsymbol{\sigma})\, U(\boldsymbol{\rho}, \boldsymbol{\sigma})\, F_0(\mathbf{r})\, \Psi_0(\mathbf{s})\, d\boldsymbol{\sigma} \tag{22}$$

with

$$\left[\nabla^2 + k^2 - \frac{8\pi^2 M_1}{h^2} V_{00}\right] F_0 = 0, \tag{23}$$

and

$$U_{\nu\nu} = \int U(\boldsymbol{\rho}, \boldsymbol{\sigma})\, |\phi_\nu(\boldsymbol{\sigma})|^2 d\boldsymbol{\sigma}. \tag{24}$$

This approximation suffers from the same defect as the method of distorted waves. If the incident nucleus spends a very long time in the neighbourhood of the struck nucleus one can no longer regard the terms which are neglected in deriving (22) as small, and the solution obtained will not satisfy the conservation of charge even approximately. Such a state of affairs will arise if the incident energy is close to a resonance value, and also if the number of particles, which can be exchanged between the two colliding nuclei, is so large that the incident energy is distributed among them all and takes a very long time to be reconcentrated on a particular one in such quantity as to result in its ejection from the system.

In attempting to obtain information in such cases we proceed as in (*a*) and neglect now all excitation processes either with or without redistribution of particles. We then write

$$\Psi = F_0(\mathbf{r})\, \psi_0(\mathbf{s}) + G_0(\boldsymbol{\rho})\, \phi_0(\boldsymbol{\sigma}). \tag{25}$$

The function Ψ can be expanded either in the form (3) or (17). The approximation (25) amounts to substituting

$$G_0(\boldsymbol{\rho})\,\phi_0(\boldsymbol{\sigma}) \quad \text{for} \quad \sum_{n\neq 0} F_n(\mathbf{r})\,\psi_n(\mathbf{s}), \tag{26}$$

and

$$F_0(\mathbf{r})\,\psi_0(\mathbf{s}) \quad \text{for} \quad \sum_{\nu\neq 0} G_\nu(\boldsymbol{\rho})\,\phi_\nu(\boldsymbol{\sigma}). \tag{27}$$

Substitution in (4) and (18) gives the simultaneous integro-differential equations

$$\left[\nabla^2 + k^2 - \frac{8\pi^2 M_1}{h^2} V_{00}\right] F_0(\mathbf{r}) = \frac{8\pi^2 M_1}{h^2} \int \psi_0^*(\mathbf{s})\,V(\mathbf{r},\mathbf{s})\,\phi_0(\boldsymbol{\sigma})\,G_0(\boldsymbol{\rho})\,d\mathbf{s}, \tag{28}$$

$$\left[\nabla_\rho^2 + \kappa^2 - \frac{8\pi^2 M_2}{h^2} U_{00}\right] G_0(\boldsymbol{\rho}) = \frac{8\pi^2 M_2}{h^2} \int \phi_0^*(\boldsymbol{\sigma})\,U(\boldsymbol{\rho},\boldsymbol{\sigma})\,\psi_0(\mathbf{s})\,F_0(\mathbf{r})\,d\boldsymbol{\sigma}. \tag{29}$$

These are the equations which we use as the basis of our calculations on capture phenomena. Using certain forms for U, V, ϕ, ψ which bear a reasonably close approximation to the actual forms, it is possible to obtain exact analytical solutions of the equations without assuming the right-hand side of (28) and (29) to be small (this is, of course, the approximation made in deriving the equations (22)). However, before we make use of the equations it is necessary to consider the interaction potentials $V(\mathbf{r},\mathbf{s})$, $U(\boldsymbol{\rho},\boldsymbol{\sigma})$ in more detail.

We have restricted ourselves to the consideration of rearrangement collisions involving two particles only, so no further loss of generality is incurred by representing the process as the replacement of one particle attached to a massive and practically unalterable core, by another. We may then take, approximately, $\mathbf{r} = \boldsymbol{\sigma}$, $\mathbf{s} = \boldsymbol{\rho}$, and write

$$V(\mathbf{r},\mathbf{s}) = V_c(\mathbf{r}) + V_i(|\mathbf{r}-\boldsymbol{\rho}|), \tag{30}$$

$$U(\boldsymbol{\rho},\boldsymbol{\sigma}) = U_c(\boldsymbol{\rho}) + V_i(|\mathbf{r}-\boldsymbol{\rho}|), \tag{31}$$

where the terms with suffix c represent interaction with the core and that with suffix i the mutual interaction between the exchanging particles. On substitution in equations (28), (29) it is found that the core terms do not vanish. If the functions G_0, ψ_0 and F_0, ϕ_0 respectively were orthogonal they would actually disappear, and we notice that the exact expressions on the right-hand side of (26), (27) are orthogonal to ψ_0, ϕ_0 respectively. We therefore discard the core terms, realizing, however, that for precise calculations with specified interactions this could not be done without closer examination. The correct substitutions to replace F_0, G_0 and satisfy the orthogonality conditions would not only remove the contribution from the core term but also alter that from the direct interaction. For our purposes this is not important as it can be regarded as merely influencing the intensity of the interaction, a quantity which we make no attempt to fix; on the contrary we aim to estimate the collision probability for wide ranges of variation of this intensity.

We then take, as basis of our calculation, the equations

$$\left[\nabla^2 + k^2 - \frac{8\pi^2 M_1}{h^2} V_{00}\right] F_0(\mathbf{r}) = \frac{8\pi^2 M_1}{h^2} \int \psi_0^*(\boldsymbol{\rho}) V_t(|\mathbf{r}-\boldsymbol{\rho}|) G_0(\boldsymbol{\rho}) \phi_0(\mathbf{r}) \, d\boldsymbol{\rho}, \quad (32)$$

$$\left[\nabla_\rho^2 + \kappa^2 - \frac{8\pi^2 M_2}{h^2} U_{00}\right] G_0(\boldsymbol{\rho}) = \frac{8\pi^2 M_2}{h^2} \int \phi_0^*(\mathbf{r}) V_t(|\mathbf{r}-\boldsymbol{\rho}|) F_0(\mathbf{r}) \psi_0(\boldsymbol{\rho}) \, d\mathbf{r}. \quad (33)$$

These equations possess the essential property that the solutions satisfy the conservation of charge. For multiplying the equations (32) and (33) by F_0^*, G_0^*, and integrating over the space of $\mathbf{r}$ and $\mathbf{s}$ respectively, gives

$$\frac{h^2}{8\pi^2 M_1} \int (F_0^* \nabla^2 F_0 - F_0 \nabla^2 F_0^*) \, d\mathbf{r} + \frac{h^2}{8\pi^2 M_2} \int (G_0^* \nabla^2 G_0 - G_0 \nabla^2 G_0^*) \, d\boldsymbol{\rho} = 0. \quad (34)$$

Employing Green's theorem gives

$$\frac{h^2}{8\pi^2 M_1} \int \mathbf{n} . (F_0^* \operatorname{grad} F_0 - F_0 \operatorname{grad} F_0^*) \, dS_1 + \frac{h^2}{8\pi^2 M_2} \int \mathbf{n} . (G_0^* \operatorname{grad} G_0 - G_0 \operatorname{grad} G_0^*) \, dS_2 = 0, \quad (35)$$

the integrals being taken over the surface of the infinite sphere in each case. Remembering that the current density corresponding to a wave function χ is

$$\frac{h}{4\pi m i} (\chi^* \operatorname{grad} \chi - \chi \operatorname{grad} \chi^*), \quad (36)$$

it is seen that equation (35) represents the vanishing of the net flux across the infinite sphere.

2—DETAILED CALCULATIONS FOR EXCITATION

We have to solve the equations (13), (14) with F_0, F_1 proper functions having the asymptotic forms

$$F_0 \sim e^{ikz} + f(\theta, \phi) r^{-1} e^{ikr}, \quad (37)$$

$$F_1 \sim f_1(\theta, \phi) r^{-1} e^{ik_1 r}. \quad (38)$$

Assuming that V_{00}, V_{11}, V_{01}, V_{10} are all spherically symmetrical, we have

$$F_0 = r^{-1}\sum_l F_0^l(r)\,P_l(\cos\theta), \quad F_1 = r^{-1}\sum_l F_1^l(r)\,P_l(\cos\theta), \tag{39}$$

where

$$\frac{d^2F_0^l}{dr^2}+\left(k^2-\frac{8\pi^2M_1}{h^2}V_{00}-\frac{l(l+1)}{r^2}\right)F_0^l = \frac{8\pi^2M_1}{h^2}V_{01}F_1^l, \tag{40}$$

$$\frac{d^2F_1^l}{dr^2}+\left(k_1^2-\frac{8\pi^2M_1}{h^2}V_{11}-\frac{l(l+1)}{r^2}\right)F_1^l = \frac{8\pi^2M_1}{h^2}V_{10}F_0^l. \tag{41}$$

The method of distorted waves then gives, when V_{01} is small, the approximate solution

$$F_0^l \sim (2l+1)\,e^{i\delta_l}\,i^l\sin(kr-\tfrac{1}{2}l\pi+\delta_l), \tag{42}$$

$$F_1^l \sim e^{ik_1r}\frac{8\pi^2M_1}{h^2}\int V_{10}(r)\,\mathscr{F}_0^l\mathscr{F}_1^l\,dr, \tag{43}$$

where $\mathscr{F}_0^l, \mathscr{F}_1^l$ are solutions of the homogeneous equations

$$\frac{d^2\mathscr{F}_0^l}{dr^2}+\left[k^2-\frac{8\pi^2M_1}{h^2}V_{00}-\frac{l(l+1)}{r^2}\right]\mathscr{F}_0^l = 0, \tag{44}$$

$$\frac{d^2\mathscr{F}_1^l}{dr^2}+\left[k_1^2-\frac{8\pi^2M_1}{h^2}V_{11}-\frac{l(l+1)}{r^2}\right]\mathscr{F}_1^l = 0, \tag{45}$$

which are finite at the origin and have the asymptotic forms

$$\left.\begin{matrix}\mathscr{F}_0^l\\ \mathscr{F}_1^l\end{matrix}\right\} \sim (2l+1)\,i^l\left\{\begin{matrix}e^{i\delta_l}\sin(kr-\tfrac{1}{2}l\pi+\delta_l),\\ e^{i\eta_l}\sin(k_1r-\tfrac{1}{2}l\pi+\eta_l).\end{matrix}\right. \tag{46}$$

Now it is also possible under certain conditions to solve the equations (13), (14) without assuming V_{01} and V_{10} small, in terms of the solutions of related homogeneous equations. Suppose

$$\frac{8\pi^2M_1}{h^2}V_{01} = \frac{8\pi^2M_1}{h^2}V_{10} = \begin{matrix}-H, & r<r_0,\\ O, & r>r_0,\end{matrix} \tag{47}$$

where H is constant. Then if V_{00}, V_{11} differ by a quantity which can also be represented by a step function we have for a region $(r<r_0)$ including the origin, in which $V_{00} = V_{11}+\dfrac{\alpha^2h^2}{8\pi^2M_1}$,

$$\frac{d^2F_0^l}{dr^2}+\left[k^2-\frac{8\pi^2M_1}{h^2}V_{00}-\frac{l(l+1)}{r^2}\right]F_0^l = HF_1^l, \tag{48}$$

$$\frac{d^2F_1^l}{dr^2}+\left[h^2-\frac{8\pi^2M_1}{h^2}V_{00}-\frac{l(l+1)}{r^2}\right]F_1^l = HF_0^l, \tag{49}$$

where $h^2 = k^2+\alpha^2$.

Let $\mathscr{F}^l(\chi, r)$ be the solution of the equation

$$\frac{d^2\mathscr{F}^l}{dr^2}+\left[\chi^2-\frac{8\pi^2 M_1}{h^2}V_{00}-\frac{l(l+1)}{r^2}\right]\mathscr{F}^l=0, \tag{50}$$

which vanishes at the origin. We try, then, as solution of the simultaneous equations (48), (49),

$$F_0^l=\mathscr{F}^l(\chi, r), \quad F_1^l=\gamma\mathscr{F}^l(\chi, r). \tag{51}$$

Then we must have

$$(k^2-\chi^2)=H\gamma, \quad \gamma(h^2-\chi^2)=H, \tag{52}$$

giving, by elimination of γ, the equation

$$(k^2-\chi^2)(h^2-\chi^2)=H^2. \tag{53}$$

If χ_1^2, χ_2^2 are the two roots of this equation, the general solution of the equations (48), (49), finite at the origin, is

$$F_0^l=A\mathscr{F}^l(\chi_1, r)+B\mathscr{F}^l(\chi_2, r), \tag{54}$$

$$F_1^l=A\gamma_1\mathscr{F}^l(\chi_1, r)+B\gamma_2\mathscr{F}^l(\chi_2, r), \tag{55}$$

with
$$\gamma_1=\frac{H}{h^2-\chi_1^2}, \quad \gamma_2=\frac{k^2-\chi_2^2}{H}.$$

In the same way the general solution can be found for a region which does not include the origin.

For detailed application of this solution we will confine ourselves to terms in which $l=0$. These are the only terms important for low velocities of impact, and for the study of resonance phenomena it is immaterial what value of l is chosen. We must now have, dropping the suffix 0,

$$F_0 \sim \sin kr+\alpha e^{ikr}, \tag{56}$$

$$F_1 \sim \beta e^{ik_1 r}, \tag{57}$$

and the contributions to the elastic and excitation cross-section are

$$\frac{4\pi}{k^2}|\alpha|^2, \quad 4\pi\frac{k_1}{k^3}|\beta|^2, \tag{58}$$

respectively.

Conservation of charge ensures that (Mott 1931)

$$4\pi\,|\beta|^2\frac{k_1}{k^3} \ngtr \frac{\pi}{k^2},$$

$$\text{i.e.}\quad \frac{k_1}{k}\,|\beta|^2 \ngtr \tfrac{1}{4}, \tag{59}$$

and

$$|\alpha|^2 \ngtr 1. \tag{60}$$

We are interested here in determining how far these quantities deviate from their maximum values under different conditions.

The forms which we take for V_{00} and V_{11} are

$$\left.\begin{aligned} \frac{8\pi^2 M_1}{h^2}V_{00} &= -C, \quad r<r_1, & \frac{8\pi^2 M_1}{h^2}V_{11} &= -C_1, \quad r<r_1', \\ &= D, \quad r_1<r<r_2, & &= D_1, \quad r_1'<r<r_2, \\ &= 0, \quad r>r_2, & &= 0, \quad r>r_2, \end{aligned}\right\} \tag{61}$$

of the familiar potential well plus potential barrier form. Although it is possible to carry the calculations through with r_0, r_1, r_1', r_2, r_2' all different this case is unnecessarily complicated, so we take $r_1' = r_1 = r_0$ leaving only two boundaries of potential discontinuity.

With these forms the method of distorted waves gives

$$|\alpha| = \sin\left[\arctan\left\{\frac{k}{\lambda}\left(\frac{\frac{\lambda}{\nu}\tan\nu r_0 - \tanh\lambda t}{1-\frac{\lambda}{\nu}\tan\nu r_0\tanh\lambda t}\right)\right\} - kr_2\right], \tag{62}$$

$$|\beta| = \frac{HKL}{k_1}(\nu_1\sin\nu r_0\cos\nu_1 r_0 - \nu\sin\nu_1 r_0\cos\nu r_0)(\nu^2-\nu_1^2)^{-1}, \tag{63}$$

where

$$\nu^2 = k^2+C, \quad \lambda^2 = D-k^2, \quad \nu_1^2 = k_1^2+C_1, \quad \lambda_1^2 = D_1-k_1^2, \quad t = r_2-r_0,$$

$$\frac{1}{K} = \frac{1}{2\lambda}\left|e^{-\lambda t}\left(1+\frac{ik}{\lambda}\right)(\lambda\sin\nu r_0-\nu\cos\nu r_0)+e^{\lambda t}\left(1-\frac{ik}{\lambda}\right)(\lambda\sin\nu r_0+\nu\cos\nu r_0)\right|. \tag{64}$$

L is derived from K by substituting ν_1, λ_1, k_1 for ν, λ, k. When

$$\lambda\sin\nu r_0+\nu\cos\nu r_0 = 0, \tag{65}$$

KL behaves like $e^{(\lambda-\lambda_1)t}$. Unless H is very small this positive exponential will result in the expression (63) being markedly in excess of the maximum value 0·5 given by (59). It is under these conditions that our exact solution is of special interest.

For the exact solution we have

$$r < r_0 = r_1 = r_1';$$

$$F_0 = a \sin \chi r + b \sin \chi_1 r, \tag{66}$$

$$F_1 = a\gamma \sin \chi r + b\gamma_1 \sin \chi_1 r, \tag{67}$$

with
$$\chi^2 = \tfrac{1}{2}(\nu^2 + \nu_1^2) + \{\tfrac{1}{4}(\nu^2 - \nu_1^2)^2 + H^2\}^{\frac{1}{2}}, \tag{68}$$

$$\chi_1^2 = \tfrac{1}{2}(\nu^2 + \nu_1^2) - \{\tfrac{1}{4}(\nu^2 - \nu_1^2)^2 + H^2\}^{\frac{1}{2}}, \tag{69}$$

$$\gamma = \frac{H}{\nu^2 - \chi_1^2}, \quad \gamma_1 = \frac{H}{\nu^2 - \chi^2}.$$

$$r_0 < r < r_2;$$

$$F_0 = ce^{-\lambda r} + de^{\lambda r}, \tag{70}$$

$$F_1 = c_1 e^{-\lambda_1 r} + d_1 e^{\lambda_1 r}. \tag{71}$$

$$r_2 < r;$$

F_0, F_1 have the asymptotic forms (56), (57) respectively.

The continuity conditions at the two boundaries give α and β as follows:

$$\frac{\alpha}{\Delta_1} = \frac{\beta}{\Delta_2} = \frac{1}{\Delta}, \tag{72}$$

where
$$\Delta = \gamma \begin{vmatrix} \sin \chi_1 r_0 & \xi_1 \\ \chi_1 \cos \chi_1 r_0 & \xi_2 \end{vmatrix} \begin{vmatrix} \sin \chi r_0 & \eta_1 \\ \chi \cos \chi r_0 & \eta_2 \end{vmatrix}$$
$$- \gamma_1 \begin{vmatrix} \sin \chi r_0 & \xi_1 \\ \chi \cos \chi r_0 & \xi_2 \end{vmatrix} \begin{vmatrix} \sin \chi_1 r_0 & \eta_1 \\ \chi_1 \cos \chi_1 r_0 & \eta_2 \end{vmatrix}; \tag{73}$$

$\Delta_1 = -\Delta$ with ζ_1, ζ_2 substituted for ξ_1, ξ_2

and
$$\Delta_2 = -k(\chi_1 \sin \chi r_0 \cos \chi_1 r_0 - \chi \cos \chi r_0 \sin \chi_1 r_0).$$

Here
$$\xi_1 = e^{ikr_2}\left(\frac{ik}{\lambda} \sinh \lambda t - \cosh \lambda t\right),$$

$$\xi_2 = e^{ikr_2}\left(\sinh \lambda t - \frac{ik}{\lambda} \cosh \lambda t\right),$$

η_1, η_2 = ξ_1, ξ_2 with λ_1, k_1 for λ, k;

$$\zeta_1 = \sin kr_2 \cosh \lambda t - \frac{k}{\lambda} \cos kr_2 \sinh \lambda t,$$

$$\zeta_2 = -\lambda \sin kr_2 \sinh \lambda t + k \cos kr_2 \cosh \lambda t. \tag{74}$$

For small values of H

$$\gamma \to \frac{H}{\nu^2-\nu_1^2}, \quad \gamma_1 \to \frac{\nu^2-\nu_1^2}{H}, \quad \chi \to \nu, \quad \chi_1 \to \nu_1,$$

$$\begin{vmatrix} \sin\chi r_0 & \xi_1 \\ \chi\cos\chi r_0 & \xi_2 \end{vmatrix} \to \frac{k}{K}, \quad \begin{vmatrix} \sin\chi_1 r_0 & \eta_1 \\ \chi_1\cos\chi_1 r_0 & \eta_2 \end{vmatrix} \to \frac{k_1}{L},$$

and the formula (74) reduces, as it must, to (63). Resonance effects arise when either of these determinants becomes exponentially small, but the exact formula (74) shows that in this case only the determinants multiplying γ_1 become small, those multiplying γ remain large and prevent the value of β exceeding the maximum given by (59). The γ term arises from the "close coupling" between the F_0 and F_1 waves and shows that, when the resonance begins to produce strong excitation the reverse process also becomes probable and reduces the chance that the system will be found in the excited state after the collision. To give quantitative expression to these effects numerical calculations have been carried out for two cases which we proceed to describe.

Numerical Results

The first case considered corresponds roughly to excitation of a light nucleus by a slow proton. The numerical values taken were:

$$\left.\begin{aligned} & r_0 = 2\times 10^{-13}\text{ cm.}, \quad t = r_2 - r_0 = 8\times 10^{-13}\text{ cm.}, \quad M_1 = 6\cdot 6\times 10^{-24}\text{ gm.}, \\ & \frac{h^2}{8\pi^2 M_1} C = 20\text{ mV} = \frac{h^2}{8\pi^2 M_1} C_1, \quad \frac{h^2}{8\pi^2 M_1} D = 2\text{ mV}, \\ & \frac{h^2}{8\pi^2 M_1}(k^2-k_1^2) = 0\cdot 05\text{ mV}. \end{aligned}\right\} \tag{75}$$

The resonance condition (65) is satisfied for incident energies of 240,000 and 290,000 e-volts, the first being due to resonance of the incoming wave, the second to the outgoing. Using (72), values of α and β as functions of the incident energy were computed for a number of values of the transition potential, $\frac{h^2}{8\pi^2 M_1} H$, ranging from $0\cdot 5\times 10^5$ to 5×10^5 e-volts. Fig. 1 (*a*) illustrates the results obtained for $\frac{k_1}{k}|\beta|^2$ given as a fraction of its maximum possible value 0·25. The dotted curve represents the form given by the

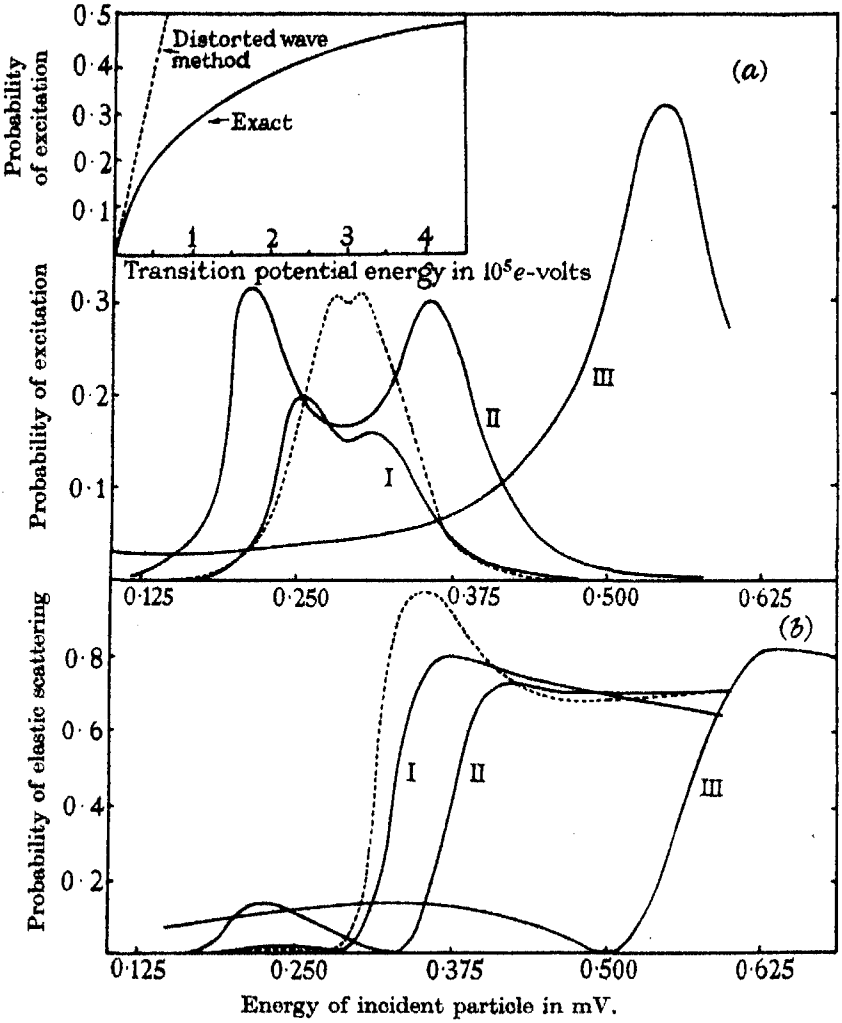


FIG. 1—Effect of strong interaction on the probability of excitation and of elastic scattering on collision of two light nuclei. (*a*) Probability of excitation as a function of the energy of the incident particles. Dotted curve—Distorted wave approximation. Curves I, II, III correspond to transition potential energies of 50,000, 125,000 and 500,000 e-volts respectively, calculated accurately. (*b*) Corresponding curves for elastic scattering. Inset—Maximum probability of excitation as a function of the transition potential energy.

method of distorted waves (formula (63)), the maximum being adjusted to a value of 0·3 for purposes of comparison; the twin maxima in this case correspond to the two close resonance points. This curve is only correct in form in the limit $H \to 0$; the effect of finite values of H is seen to consist in a separation of the maxima by an increasing amount accompanied by an increasing breadth of the resonance range, this broadening effect being very marked. For $\frac{h^2 H}{8\pi^2 M_1} = 1{\cdot}2 \times 10^5$ e-volts the probability of excitation remains within 0·1 of the maximum possible value over a range of incident energies as large as 250,000 e-volts. Values of $\frac{h^2 H}{8\pi^2 M_1}$ comparable with 10^6 e-volts are not at all impossible, so we must expect to find this broadening in actual practice. The actual form of the wings of the resonance curves is not to be relied upon, for we have taken a rectangular potential barrier which probably gives rather different results from the actual Coulomb barrier in this respect. The inset in fig. 1 gives the value of $\frac{k_1}{k} |\beta|^2$ at the resonance maximum as a function of H. It never attains the maximum possible value of 0·25 but reaches a smaller maximum of about one-half this value and then falls off very slowly for larger values of H.

The corresponding behaviour of $|\alpha|^2$, which determines the elastic scattering, is given in fig. 1 (*b*). The sudden change in this quantity near the resonance maximum is very marked. This effect was predicted by Mott (1931) for the case $H = 0$, but it is seen that the effect persists at the modified resonance maximum which appears for finite values of H. It is apparent that when resonance can occur the influence of the possibility of excitation on the elastic scattering is very marked, and it is therefore impossible to argue about nuclear fields from the observed elastic scattering without taking this into account.

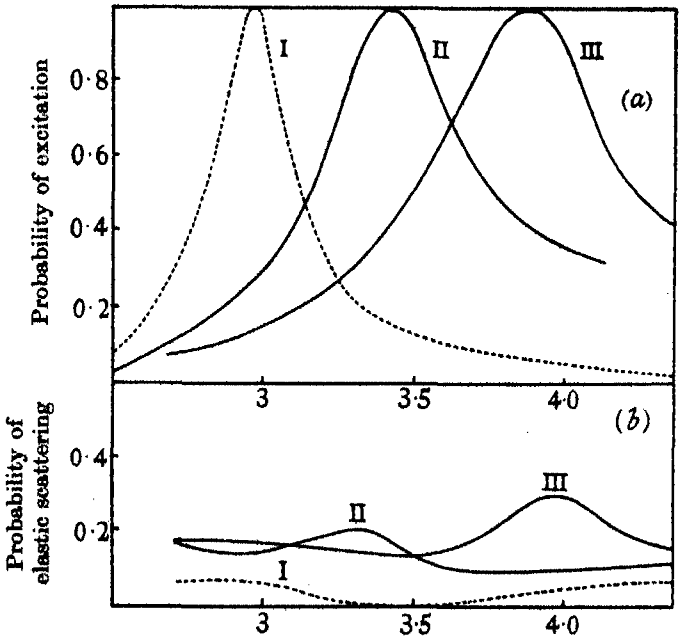
An important result which emerges is that the maximum yield occurs at an incident energy which depends on all three interaction terms V_{00}, V_{11}, V_{01} and not on just V_{00} or V_{11} alone. This must be taken into account even for quite small values of V_{01}. Analogous results will be found in § 3 for rearrangement collisions where it is possible to interpret the effect as involving resonance with a two- or multi-body system instead of just a one-body effect.

The second case which we have considered corresponds to excitation of a nucleus by a charged particle which possesses sufficient energy to enter over the potential barrier but loses so much energy in effecting excitation that it must pass out through the barrier. It is supposed that the conditions are such that resonance may occur for certain energies of the outgoing particles. Numerical values assumed were

$$\left.\begin{array}{l} r_0 = 3{\cdot}5\times 10^{-13}\,\text{cm.},\quad t = r_2 - r_0 = 3\times 10^{-13}\,\text{cm.},\quad M_1 = 6{\cdot}6\times 10^{-28}\,\text{gm.}, \\ \dfrac{h_2}{8\pi^2 M_1}C = 16\,\text{mV} = \dfrac{h_2}{8\pi^2 M_1}C_1,\quad \dfrac{h_2}{8\pi^2 M_1}D = 2{\cdot}5\,\text{mV} = \dfrac{h_2}{8\pi^2 M_1}D_1, \\ \dfrac{h^2}{8\pi^2 M_1}(k^2 - k_1^2) = 2\,\text{mV}. \end{array}\right\} \tag{76}$$

These values correspond roughly to the hypothetical excitation of a lithium nucleus by an α-particle. The resonance condition (63) is satisfied by the constants of the outgoing wave when the energy of the incident particle is 3 mV.

Fig. 2 illustrates results for the probability of excitation $\dfrac{4k_1}{k^3}|\beta|^2$. The dotted curve I is the form of the excitation curve in the limit $H\to 0$ given



Energy of incident α-particle in mV.

FIG. 2—As for fig. 1, but with different choice of parameters. (*a*) Excitation; (*b*) elastic scattering. Dotted curves I—Distorted wave approximation. Curves II, III correspond to transition potential energies of 1·6 and 2·5 mV respectively, calculated accurately.

by the distorted wave formula (63) with the maximum adjusted to unity. Curves II and III correspond to values of $\frac{h^2}{8\pi^2 M_1} H$ of 1·6 and 2·5 mV respectively. The shift of the energy of maximum yield with increasing H is very marked, as also is the broadening. It is noteworthy that even for the larger value of H the maximum probability of unity is attained despite the damping effect of the strong interaction.

In the limit $H \to 0$ the elastic scattering exhibits no resonance effects in this energy range as these arise, in this approximation, from the outgoing wave of diminished energy only; but when H is finite the back coupling between the elastic and inelastic collisions introduces the resonance effects shown clearly in fig. 2 (*b*), curves II and III.

3—DETAILED CALCULATIONS FOR CAPTURE

We have to solve the equations (37), (38) with F_0, G_0 proper functions having the asymptotic forms

$$F_0 \sim e^{ikz} + f(\theta, \phi)\, r^{-1} e^{ikr},$$

$$G_0 \sim g(\vartheta, \psi)\, \rho^{-1} e^{i\kappa\rho}.$$

Assuming the functions $V_{00}, U_{00}, \psi_0, \phi_0$ to be spherically symmetrical, we have

$$F_0 = r^{-1}\, \Sigma_l F_0^l(r)\, P_l(\cos\theta), \quad G_0 = \rho^{-1}\, \Sigma_l G_0^l(\rho)\, P_l(\cos\theta), \tag{77}$$

where

$$\frac{d^2 F_0^l}{dr^2} + \left(k^2 - \frac{l(l+1)}{r^2} - \frac{8\pi^2 M_1}{h^2} V_{00}\right) F_0^l = \frac{32\pi^3 M_1}{h^2(2l+1)}\, r\phi_0(r) \times \int_0^\infty \psi_0^*(\rho)\, V_i^l(r, \rho)\, G_0^l(\rho)\, \rho\, d\rho, \tag{78}$$

$$\frac{d^2 G_0^l}{dr^2} + \left(\kappa^2 - \frac{l(l+1)}{r^2} - \frac{8\pi^2 M_2}{h^2} U_{00}\right) G_0^l = \frac{32\pi^3 M_2}{h^2(2l+1)}\, \rho\psi_0(\rho) \times \int_0^\infty \phi_0^*(r)\, V_i^l(r, \rho)\, F_0^l(r)\, r\, dr, \tag{79}$$

and

$$V_i(|\mathbf{r} - \boldsymbol{\rho}|) = \Sigma V_i^n P_n\left(\frac{\mathbf{r}\cdot\boldsymbol{\rho}}{r\rho}\right). \tag{80}$$

As for excitation we confine ourselves to the case $l = 0$ and drop the index 0, so

$$F_0 \sim \sin kr + \alpha e^{ikr},$$

$$G_0 \sim \beta e^{i\kappa r},$$

where, again,
$$\frac{\kappa}{k}|\beta|^2 \not> \tfrac{1}{4}, \tag{81}$$

$$|\alpha|^2 \not> 1. \tag{82}$$

The method of distorted waves gives

$$F_0 \sim e^{i\delta_0} \sin(kr+\delta_0), \text{ as before,} \tag{83}$$

$$G_0 \sim e^{i\kappa\rho} \frac{32\pi^3 M_2}{\kappa h^2} \int_0^\infty \int_0^\infty \mathscr{F}_0(r)\, \psi_0(\rho)\, V_i^0(r,\rho) \mathscr{G}_0(\rho) \phi_0^*(r)\, dr\, d\rho, \tag{84}$$

where $\mathscr{F}_0(r), \mathscr{G}_0(\rho)$ are the solutions of the homogeneous equations obtained by placing $V_i^0(r,\rho) = 0$ in (75), (76), which vanish at the origin and have the asymptotic forms

$$\mathscr{F}_0(r) \sim \sin(kr+\delta_0)\, e^{i\delta_0}, \tag{85}$$

$$\mathscr{G}_0(r) \sim e^{i\eta_0} \sin(\kappa r+\eta_0). \tag{86}$$

To obtain an exact solution we make the following idealizations of the functional form of the various quantities involved:

$$\begin{aligned}
\frac{8\pi^2 M_1}{h^2} V_{00} &= -C, \quad r<r_0, & \frac{8\pi^2 M_2}{h^2} U_{00} &= -C', \quad \rho<r_0,\\
&= D, \quad r_0<r<r_1, & &= D', \quad r_0<\rho<r_1,\\
&= 0, \quad r>r_1, & &= 0, \quad \rho>r_1.
\end{aligned}$$

$$\begin{aligned}
\psi_0 &= \left(\frac{3}{4\pi r_0^3}\right)^{\frac{1}{2}}, \quad r<r_0, \quad \phi_0 = \psi_0 \text{ with } \rho \text{ for } r,\\
&= 0, \quad r>r_0,
\end{aligned}$$

$$\frac{8\pi^2 M_1}{h^2} V_i(|r-\rho|) = \frac{A}{|\mathbf{r}-\boldsymbol{\rho}|}, \quad r,\rho<r_0. \tag{87}$$

With these forms the equations can be solved exactly by an extension of the method used by Morse and Allis (1933) for the solution of a similar pair of equations occurring in the theory of the effect of exchange on the elastic scattering of electrons by atoms. We first note that the distorted wave formulae (83), (84) give now

$|\alpha|$, as in (62),

$$|\beta| = \frac{3A}{r_0^3} \frac{KL'}{\kappa\nu^2} \frac{M_2}{M_1} \left[\frac{1}{2}\left\{\frac{\sin(\nu'-\nu)r_0}{\nu'-\nu} - \frac{\sin(\nu'+\nu)r_0}{\nu'+\nu}\right\} - \nu\cos\nu r_0\left(-\frac{r_0}{\nu'}\cos\nu' r_0 + \frac{\sin\nu' r_0}{\nu'^2}\right)\right], \tag{88}$$

where

$$\nu^2 = k^2 + C, \quad \nu'^2 = \kappa^2 + C', \quad K \text{ as in (64)},$$

$$\begin{aligned} L' = \lambda'^{-1}\Big|&(\lambda'\sin\nu' r_0\cos\lambda' r_0 - \nu'\cos\nu' r_0\sin\lambda' r_0)\left(\cos\lambda' r_1 + \frac{i\lambda'}{\kappa}\sin\lambda' r_1\right) \\ &+ (\lambda'\sin\nu' r_0\sin\lambda' r_0 + \nu'\cos\nu' r_0\cos\lambda' r_0)\left(\sin\lambda' r_1 - \frac{i\lambda'}{\kappa}\cos\lambda' r_1\right)\Big|, \end{aligned}$$

$$\lambda'^2 = \kappa^2 - D'. \tag{89}$$

For the exact solution we proceed as follows:

$r < r_0$, $\rho < r_0$. We try as solution

$$F = a_1\sin\chi_1 r + a_2\sin\chi_2 r + a_3\sin\chi_3 r + a_4\sin\chi_4 r, \tag{90}$$

$$G = a_1'\sin\chi_1\rho + a_2'\sin\chi_2\rho + a_3'\sin\chi_3\rho + a_4'\sin\chi_4\rho. \tag{91}$$

Then we must have

$$\sum_{s=1}^{4} a_s(-\chi_s^2 + \nu^2)\sin\chi_s r = \frac{3A}{r_0^3}\left[r\,\Sigma_s a_s'\frac{\cos\chi_s r_0}{\chi_s} - \Sigma a_s'\frac{\sin\chi_s r}{\chi_s^2}\right], \tag{92}$$

$$\sum_{s=1}^{4} a_s'(-\chi_s^2 + \nu'^2)\sin\chi_s\rho = \frac{3A}{r_0^3}\frac{M_2}{M_1}\left[\rho\,\Sigma_s a_s\frac{\cos\chi_s r_0}{\chi_s} - \Sigma a_s\frac{\sin\chi_s\rho}{\chi_s^2}\right], \tag{93}$$

so

$$\Sigma_s a_s\frac{\cos\chi_s r_0}{\chi_s} = 0, \tag{94}$$

$$\Sigma_s a_s'\frac{\cos\chi_s r_0}{\chi_s} = 0, \tag{95}$$

and

$$a_s(-\chi_s^2 + \nu^2) = -\frac{3A}{r_0^3}\frac{a_s'}{\chi_s^2}, \tag{96}$$

$$a_s'(-\chi_s^2 + \nu'^2) = -\frac{3A}{r_0^3}\frac{M_2 a_s}{M_1\chi_s^2}. \tag{97}$$

Eliminating the ratio a_s/a_s' gives the biquadratic equation

$$\chi_s^4(-\chi_s^2 + \nu^2)(-\chi_s^2 + \nu'^2) = \frac{9A^2}{r_0^6}\frac{M_2}{M_1} \tag{98}$$

for χ. The four roots give the values of the ratios a_s/a_s', so providing, in conjunction with (92) and (93), six equations to determine the eight quantities a_1, a_2, a_3, a_4; a_1', a_2', a_3', a_4'. The continuity conditions to be satisfied at the boundaries give the remaining two relations which make the solution completely determinate:

$r_0 < r < r_1$; $r_0 < \rho < r_1$;

$$F_0 = ce^{-\lambda r} + de^{\lambda r},$$

$$G_0 = c'\sin\lambda'\rho + d'\cos\lambda'\rho.$$

(The outgoing particle is supposed to have sufficient energy to pass over the barrier.)

$$r > r_1;\quad \rho > r_1;\qquad F_0 = \sin kr + \alpha e^{ikr},\quad G_0 = \beta e^{ik\rho}.$$

Solving the equations expressing the continuity conditions gives

$$\frac{\alpha}{\Delta_1} = \frac{\beta}{\Delta_2} = \frac{1}{\Delta}, \tag{99}$$

where

$$\Delta = \begin{vmatrix} \chi_1^{-1}\cos\chi_1 r_0, & \ldots, & \ldots, & \chi_4^{-1}\cos\chi_4 r_0, & 0, & 0 \\ \sin\chi_1 r_0, & \ldots, & \ldots, & \sin\chi_4 r_0, & 0, & -B_1 \\ \chi_1\cos\chi_1 r_0, & \ldots, & \ldots, & \chi_4\cos\chi_4 r_0, & 0, & -B_2 \\ p_1\chi_1^{-1}\cos\chi_1 r_0, & \ldots, & \ldots, & p_4\chi_4^{-1}\cos\chi_4 r_0, & 0, & 0 \\ p_1\sin\chi_1 r_0, & \ldots, & \ldots, & p_4\sin\chi_4 r_0, & -A_1, & 0 \\ p_1\chi_1\cos\chi_1 r_0, & \ldots, & \ldots, & p_4\chi_4\cos\chi_4 r_0, & -A_2, & 0 \end{vmatrix} \tag{100}$$

$\Delta_1 = \Delta$ with the fifth column replaced by the sixth and the sixth by $0, 0, 0, 0, C_1, C_2$;

$\Delta_2 = \Delta$ with the sixth column replaced by $0, 0, 0, 0, C_1, C_2$.

$$\left.\begin{aligned} A_1 &= \cosh\lambda t - \frac{ik}{\lambda}\sinh\lambda t, \\ A_2 &= \lambda\left(-\sinh\lambda t + \frac{ik}{\lambda}\cosh\lambda t\right), \\ B_1 &= \cos\lambda' t - \frac{ik}{\lambda'}\sin\lambda' t, \\ B_2 &= \lambda'\left(\sin\lambda' t + \frac{ik}{\lambda'}\cos\lambda' t\right), \\ C_1 &= \cosh\lambda t\sin kr_1 - \frac{k}{\lambda}\sinh\lambda t\cos kr_1, \\ C_2 &= \lambda\left(-\sinh\lambda t\sin kr_1 + \frac{k}{\lambda}\cosh\lambda t\cos kr_1\right), \\ p_s &= a_s/a'_s = -\frac{3A}{r_0^3}\chi_s^{-2}(-\chi_s^2 + \nu^2)^{-1}, \\ & r_1 - r_0 = t. \end{aligned}\right\} \tag{101}$$

This formula can be shown to reduce to (90) when $A \to 0$.

Numerical Evaluation

The formula (99) was worked out in detail for the following values of the constants involved:

$$r_0 = 3\times 10^{-13}\text{ cm.},\quad t = r_1 - r_0 = 3\times 10^{-13}\text{ cm.},$$

$$M_1 = 4{\cdot}4\times 10^{-24}\text{ g.},\quad M_2/M_1 = \tfrac{1}{3},$$

$$\frac{h^2}{8\pi^2 M_1}C = 51\text{ mV},\qquad \frac{h^2}{8\pi^2 M_1}C' = 9\text{ mV},$$

$$\frac{h^2}{8\pi^2 M_1}D = 3{\cdot}5\text{ mV},\qquad D' = 0.$$

These values were chosen as being roughly representative of the reaction

$${}_4\mathrm{Be}^9 + {}_2\mathrm{He}^4 \rightarrow {}_6\mathrm{C}^{12} + {}_0n^1,$$

and lead to a resonance level at $2{\cdot}5\times 10^6$ mV for $A = 0$. Calculations were then carried out for various values of A and the results are illustrated in fig. 3.

To provide some convenient measure of the magnitude of A we give the value of $\dfrac{8\pi^2 M_1}{h^2}\dfrac{A}{r} = B$ at $r = 3\times 10^{-13}$ cm. in mV instead of that of A alone. In this way we relate the curves to the effective interaction between the neutron and α-particle, though this need not be the interaction between them when uninfluenced by other particles; large values of B correspond in our simplified model to a large deviation from the distorted wave approximation and this may arise from a variety of causes.

Fig. 3 (a) refers to the probability of capture as a function of the energy of the incident particle; the dotted curve is that given by the distorted wave method for $B = 88$ mV. Comparison with the analogous curves for excitation (fig. 1 (a)) shows that the broadening of the resonance by the strong interaction is not nearly so marked in this case. The reason for this appears to be that for the values of A concerned a second resonance effect is superimposed as revealed in the inset of fig. 3 (a). In this figure the maximum capture probability (as a function of energy) is illustrated as a function of A and is seen to be greater initially than that given by the distorted wave approximation (the dotted curve in the figure). In terms of the formula (99) the effect arises from the value of $\chi_1 r_0$ being in the neighbourhood of $\frac{3}{2}\pi$ for these values of A, while at the same time the modified resonance condition relating the quantities χ, λ, λ', k and K which produces the maximum in the probability-energy curve is also satisfied. This shows clearly that the resonance phenomena cannot be regarded in terms of the motion of a particle in a potential field but are determined by the virtual energy levels of the complex (core + incident particle + replaced particle) formed on collision. This is also clearly shown by the dependence of the position of the maximum in the probability-energy curve on the value of A.

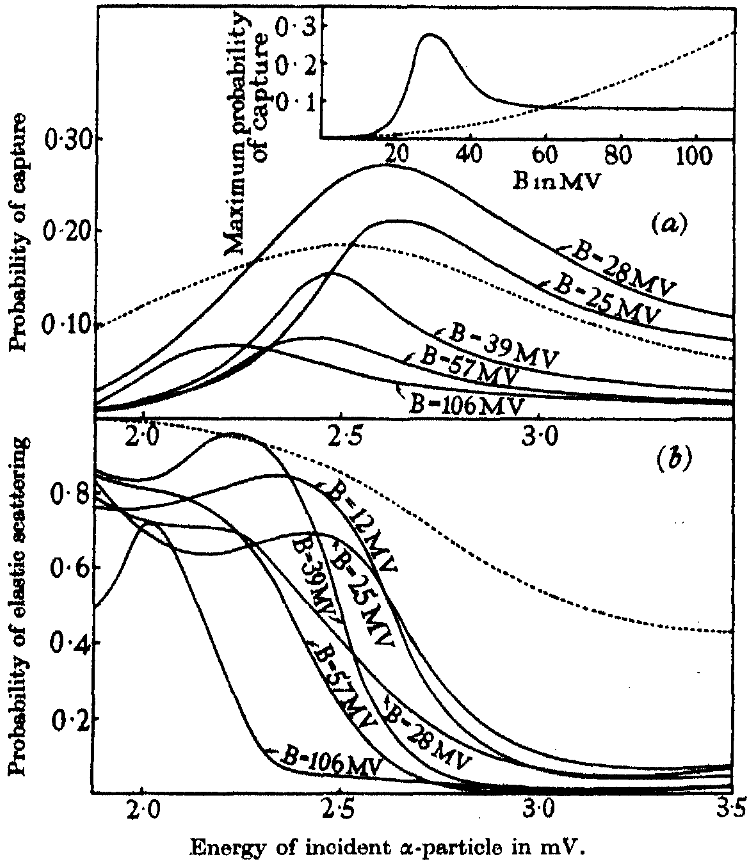


FIG. 3—Effect of strong interaction on the probability of capture of a charged particle by a light nucleus with neutron emission. (*a*) Probability of capture as a function of the energy of the incident particles for various values of the interaction parameter B. (*b*) Corresponding curves for elastic scattering. Inset. Maximum probability of capture as a function of the interaction parameter. Dotted curve—Distorted wave approximation. Full curve—Accurate calculation.

An interesting feature of the results is that the probability of capture is always considerably less than the upper limit of 0·25, a result predicted by Mott (1931) on general grounds.

The elastic scattering is represented in fig. 3 (*b*). The general behaviour is exactly similar to that illustrated in fig. 1 (*b*) and the remarks made in § 2 again apply; the sudden drop in the scattering as the energy increases through the position of the resonance maximum is a noteworthy feature of both cases.

4—CAPTURE OF LOW-VELOCITY NEUTRONS

We now consider the case where the incident particle is moving very slowly and is uncharged while the ejected charged particle receives sufficient energy by the rearrangement to pass out over the potential barrier. The low-velocity limit of the elastic (Q_0) and capture (Q_c) cross-sections can be written down somewhat more explicitly than for higher velocities. They are

$$\operatorname*{Lt}_{k \to 0} Q_0 = \left|\frac{G}{H}\right|^2 = |\gamma|^2, \text{ say,} \tag{102}$$

$$\operatorname*{Lt}_{k \to 0} Q_c = \frac{\kappa}{k}\left|\frac{I}{H}\right|^2 = \frac{\kappa}{k}|\delta|^2, \text{ say,} \tag{103}$$

where

$H = JB_1 - NB_2,$

$G = (Jr_0 - P)B_1 - (Nr_0 - Q)B_2,$

$I = \Sigma_s(-1)^s i_s \chi_s \tan \chi_s r_0,$

$J = \Sigma_s(-1)^s i_s p_s \chi_s^2,$

$N = \Sigma_s(-1)^s n_s \chi_s \tan \chi_s r_0,$

$P = \Sigma_s(-1)^s i_s p_s \chi_s \tan \chi_s r_0,$

$Q = \Sigma_s(-1)^s q_s p_s \chi_s \tan \chi_s r_0,$

$i_1' = \chi_2^2(p_3 - p_4) + \chi_3^2(p_4 - p_2) + \chi_4^2(p_2 - p_3)$ and i_2, i_3, i_4 follow in cyclic order,

$n_1 = p_4\chi_4^2(p_3 - p_2) + p_3\chi_3^2(p_2 - p_4) + p_2\chi_2^2(p_4 - p_3),$

$q_1 = \chi_2 \tan \chi_2 r_0(p_3 - p_4) + \chi_3 \tan \chi_3 r_0(p_4 - p_1) + \chi_4 \tan \chi_4 r_0(p_2 - p_3).$

$p_1, \ldots, p_4$; B_1, B_2 are as defined in (103).

Near the low-velocity limit Q_c varies inversely as the velocity and so becomes very great. What interests us here is the behaviour of the multiplying factor $|\delta|^2$ near the limit and also that of the elastic scattering $|\gamma|^2$

Numerical calculations were first carried out for the following values of the constants, very roughly representative of the reaction:

$$ {}_5\mathrm{B}^{10} + {}_0n^1 \rightarrow {}_3\mathrm{Li}^7 + {}_2\mathrm{He}^4; \qquad (104)$$

$$M_1 = 1{\cdot}5 \times 10^{-24}\ \mathrm{g.},\quad M_2 = 2{\cdot}8M_1,\quad t = 2 \times 10^{-13}\ \mathrm{cm.},$$

$$\frac{h^2}{8\pi^2 M_1}C = 23\ \mathrm{mV},\quad \frac{h^2}{8\pi^2 M_2}C' = 6{\cdot}5\ \mathrm{mV},$$

$$\frac{h^2}{8\pi^2 M_2}D = 2{\cdot}2\ \mathrm{mV},\quad \frac{h^2}{8\pi^2 M_2}\kappa^2 = 2{\cdot}7\ \mathrm{mV},$$

and various values of A. The results are illustrated in fig. 4 for three values of r_0.

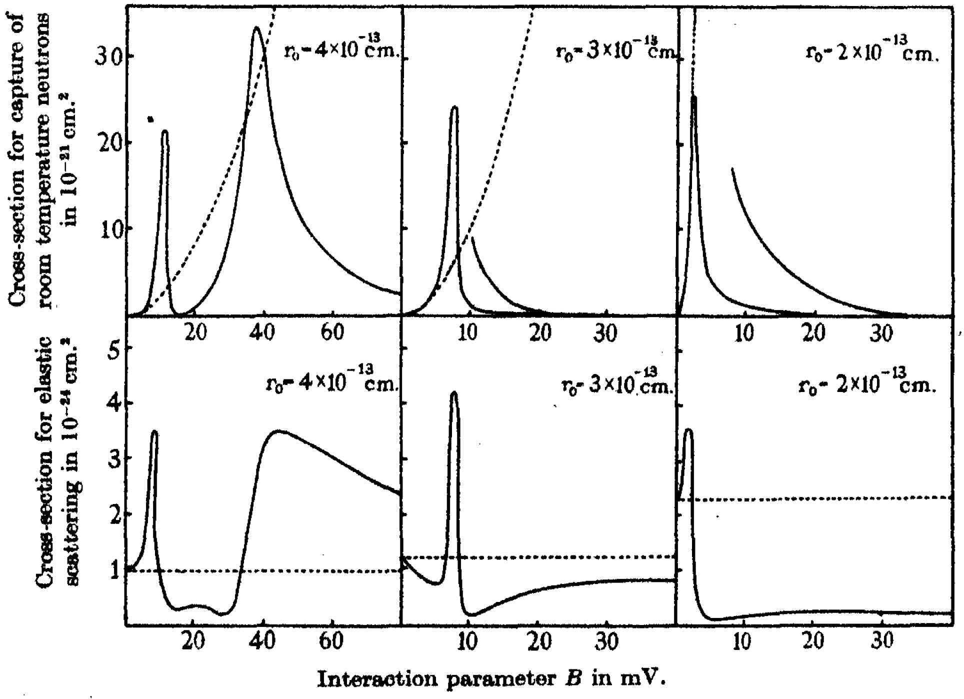


FIG. 4—Cross-sections for capture and elastic scattering of slow neutrons by light nuclei (accompanied by α-particle emission), as functions of the interaction parameter B. Dotted curves refer to the distorted wave approximation, full-line curves to the accurate calculation.

To facilitate comparison with experiment the capture cross-sections are given for neutrons with kinetic energies $\frac{3}{2}kT$, where k is Boltzmann's constant and T is taken as 300° K., i.e. for "room temperature" neutrons.

The dotted curves given by the method of distorted waves are of specially simple form. For capture they are parabolic, while for elastic scattering the cross-section is independent of A. The full-line curves, representing the results of our more accurate method, are by no means so simple, maxima appearing in both cross-sections for certain values of A while at other values the capture cross-section is comparatively small. These resonance effects (which are to be compared with that illustrated in the inset of fig. 3) can only be considered as arising from the multi-body nature of the phenomena, depending as they do on the interaction parameter A. The oscillatory nature of the probability of an inelastic collision as a function of a parameter specifying the intensity of interaction, when this is strong, has been discussed qualitatively by Mott and Massey (1933, p. 104).

The capture cross-section for the reaction (104) for room-temperature neutrons has been given by Chadwick and Goldhaber (1933) as about 4×10^{-21} cm.^{2}, and that for the reaction

$$_3\mathrm{Li}^6 + {}_0n^1 \rightarrow {}_2\mathrm{He}^4 + {}_1\mathrm{H}^3$$

as 4×10^{-22} cm.^{2}. These values are comparable with those of fig. 4, the first being in the neighbourhood of a resonance point, the second elsewhere.

In fig. 5 the probability of capture is given as a function of the average energy of interaction $\frac{h^2}{8\pi^2 M_1} C$ between the $_0n^1$ and $_5\mathrm{B}^{10}$ nuclei. A resonance effect of half-width 3 mV is clearly marked, but both cross-sections remain finite at the maximum whereas the distorted wave method gives infinite values where $Cr_0^2 = (2n+1)^2 \frac{\pi^2}{4}$. A similar resonance effect will appear if the probability is evaluated as a function of the energy of the incident neutrons, but the half-width will be much greater than that observed in experiments on radiative capture of slow neutrons by nuclei since we have only considered a two-body problem. An essential feature of the capture of slow neutrons by the heavier nuclei is that the nucleus capturing the neutron contains a large number of closely connected particles and the lifetime of the collision complex becomes very long.

The effect of varying the width t of the potential barrier acting on the outgoing α-particle was examined and found to consist merely in an alteration of the magnitudes of the cross-sections but not of their form.

5—CONCLUDING REMARKS

It is clear from the detailed calculations described in §§ 2, 3, 4 that the quantum theory of collisions when applied to nuclear phenomena must be worked out to a higher degree of approximation than customary in application to other types of collision. The results we have obtained do contain some of the essential features of Bohr's theory, but cannot be applied

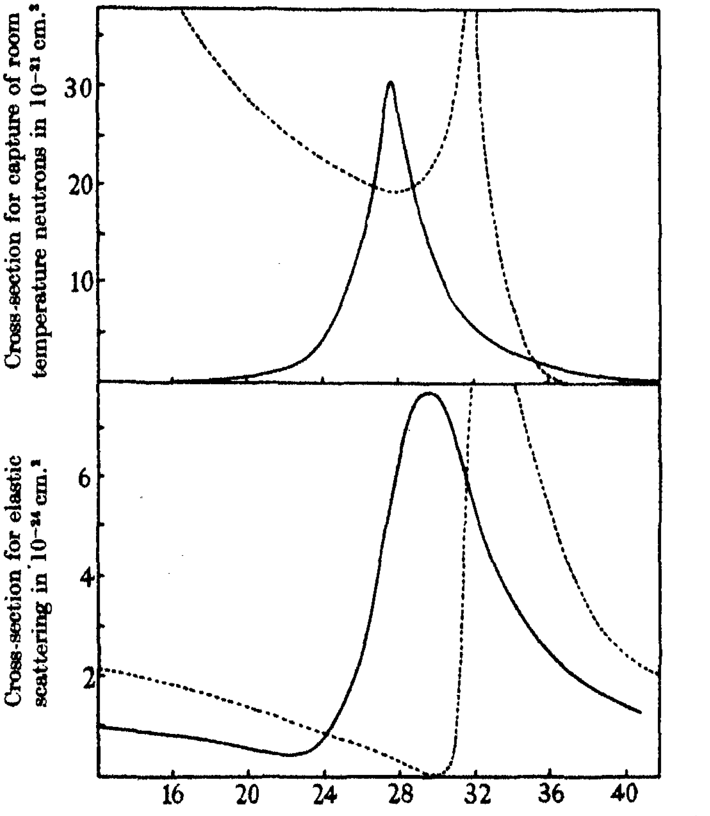


FIG. 5—Cross-sections for capture and elastic scattering of slow neutrons by light nuclei (accompanied by α-particle emission), as functions of the average interaction energy between the neutron and the nucleus. The parameters are adjusted to correspond roughly to the reaction

$$_0n^1 + {_5B^{10}} \rightarrow {_2He^4} + {_3Li^7}.$$

Dotted curves refer to the distorted wave approximation, full-line curves to the accurate calculation.

to quantitative comparison with experiment as we have only dealt with problems of a two-body (or rather three-body if the nuclear core is included) nature and the experimental results refer to multi-body phenomena. The qualitative features are reproduced and it is hoped to extend the calculations to three-body (or four-body including the nuclear core) systems. It would then be possible to investigate the effect of an increase in the number of particles concerned in the form of the resonance effects.

We wish to express our indebtedness to the Government Grant Committee of the Royal Society for providing us with funds to employ the professional services of Mr. R. Campbell, B.Sc., for performing the numerical calculations.

SUMMARY

The quantum theory of collisions must be worked out to a higher degree of accuracy than hitherto if it is to be capable of describing the resonance phenomena which are a striking feature of collisions between nuclei. In particular it is essential that the assumption of predominance of elastic scattering be not involved. In this paper calculations have been carried out for excitation and capture phenomena without this assumption. The complication is so great that the work in this paper is confined to cases which are of a two-body nature (or three-body if the nuclear core is included), but the results provide qualitative and, in some cases, some quantitative information concerning the resonance effects to be expected. They also show that the resonance phenomena must be related to energy levels of the semi-stable complex formed in the intermediate stage of the collision and cannot be understood in terms of energy levels of a single particle moving in a static potential field.

REFERENCES

Bethe 1935 *Phys. Rev.* **47**, 747.
Bethe and Placzek 1937 *Phys. Rev.* **51**, 450.
Bohr 1936 *Nature, Lond.*, **137**, 344.
Breit and Wigner 1936 *Phys. Rev.* **49**, 519.
Chadwick and Goldhaber 1933 *Proc. Camb. Phil. Soc.* **31**, 612.
Elsasser and Perrin 1935 *C.R. Acad. Sci., Paris*, **200**, 450.
Feather 1937 "Introduction to Nuclear Physics." Camb. Univ. Press.
Morse and Allis 1933 *Phys. Rev.* **44**, 269.
Mott 1931 *Proc. Roy. Soc.* A, **133**, 328.
Mott and Massey 1933 "The Theory of Atomic Collisions." Oxford Univ. Press.

The Inhibition by Nitric Oxide of the Decomposition of Ethane

BY L. A. K. STAVELEY

(*Communicated by C. N. Hinshelwood, F.R.S.—Received 8 July 1937*)

INTRODUCTION

The inhibiting effect of nitric oxide on the thermal decomposition of certain organic compounds has been attributed to the ability of this substance to suppress reaction chains by removing free radicals from the system (Staveley and Hinshelwood 1936*a*, *b*, *c*; 1937). The inhibition has been observed with various ethers and with propionic and butyric aldehydes, but the decompositions of acetaldehyde, chloral, methyl alcohol and ketones are exempt. It has been inferred that in these latter reactions either free radicals are not produced, or else that they are incapable of effecting the further decomposition of their parent substance.

The thermal decomposition of ethane has been found to be considerably retarded by nitric oxide, and in view of the comparative simplicity of the ethane molecule it seemed profitable to examine the inhibited reaction in some detail.

The normal decomposition has been extensively investigated; a full list of references will be found in a paper by Sachsse (1936*a*). In a silica bulb the reaction is predominantly homogeneous. The primary products are ethylene and hydrogen, but at temperatures at which the rate of the reaction can be conveniently measured the decomposition is incomplete. The ethylene produced slowly undergoes further changes with the formation of methane and higher hydrocarbons, but Sachsse (1936*a*) has shown that in the early stages of the reaction the pressure increase gives an accurate measure of the extent to which the primary decomposition has proceeded. These reasons, and the facts that ethylene reacts with nitric oxide and moreover itself somewhat retards the decomposition, necessitated the determination of initial rates as giving the only true measure of the reaction velocity.

EXPERIMENTAL METHOD

The apparatus used was essentially the same as that described in previous papers, with the exception that when the pressure change was fairly slow it was followed by means of a Bourdon gauge. This was made of very thin soft glass and mounted in a cylindrical glass tube. It was connected to a mercury manometer and the reaction vessel, and fitted with a long pointer damped in butyl phthalate. When ethane was admitted into the reaction bulb, dry air was let into the space outside the gauge at the same rate, so that at the beginning of the run the pointer was almost in its equilibrium position. Its movement was then followed by means of a micrometer eyepiece. The relation between the pressure difference between the inside and outside of the gauge and the displacement of the pointer proved to be an exactly linear one up to pressure differences of about 3 mm. A pressure change of 1 mm. produced a displacement of the image of the pointer in the eyepiece of 15 scale divisions. Accurate values for the initial rate could be obtained in this way since it was only necessary to follow the reaction over the first 2 or 3 % of the decomposition, in which range the plot of pressure increase against time was almost if not quite linear. The volume of the gauge was small compared with that of the reaction bulb. When the pressure change was too rapid to allow the Bourdon gauge to be used the decomposition was followed with a mercury manometer, with the aid, if necessary, of a cathetometer. A silica reaction-bulb was used.

The temperature of the furnace in which the reaction vessel was heated was measured by means of a platinum platinum-rhodium thermocouple, the cold junction being kept in melting ice. The very high temperature coefficient of the inhibited decomposition made the minimization of temperature errors essential, and this was effected by observing the pyrometer needle through a micrometer eyepiece so that any temperature fluctuations could be checked before they became serious. Delicate control of the current heating the furnace was secured by connecting an adjustable resistance in parallel with the main rheostat in the circuit. In this way the temperature during a run could be kept constant to within ± 0·3° C.

Commercial ethane of at least 97 % purity was condensed in liquid air and the residual gases pumped off. It was then partially evaporated, recondensed, and the residue again pumped off. This process was repeated several times until the vapour pressure of the ethane at the temperature of liquid air was less than 0·5 mm. About two-thirds of the liquid was evaporated into reservoirs. In some preparations the ethane before being condensed was bubbled slowly through alkaline pyrogallol to remove any traces of oxygen, but all samples gave the same results. A comparison of the rate of decomposition with that given by Sachsse indicated a difference in the two temperature scales of about 1° C.

EXPERIMENTAL RESULTS

The effect of nitric oxide on the initial rate of decomposition at 620° C. with an ethane pressure of 150 mm. is shown in fig. 1 and Table I. A

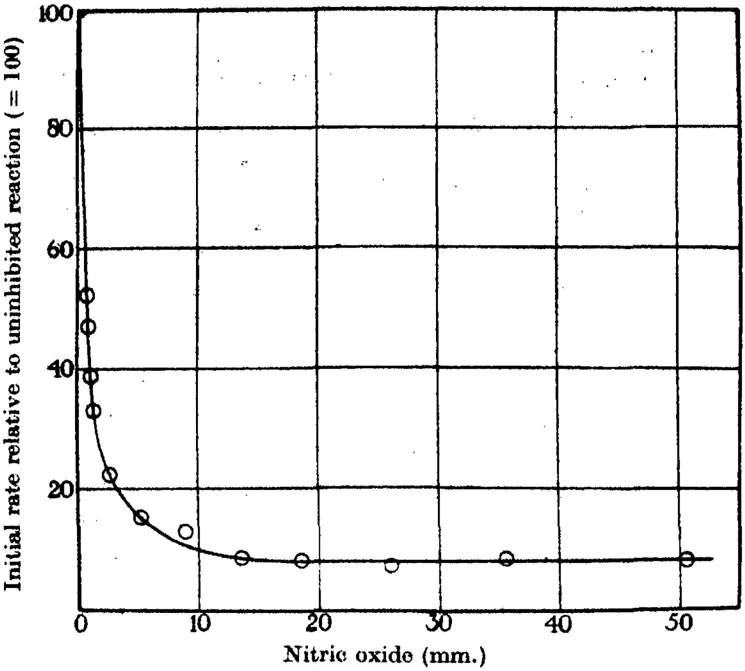


FIG. 1—Inhibition of the decomposition of ethane.

TABLE I—VARIATION OF THE INITIAL RATE OF DECOMPOSITION OF ETHANE WITH THE NITRIC OXIDE PRESSURE

Nitric oxide pressure (mm.)	Velocity constant (sec.$^{-1}$ × 10^5)
0	58·5
0·65	30·6
0·79	27·6
0·97	22·8
1·25	19·3
2·52	13·1
5·1	8·9
8·9	7·55
13·5	4·95
18·5	4·75
26·0	4·25
35·5	5·0
50·5	4·8

steady limit is reached, but the quantity of nitric oxide required to produce the maximum inhibition is somewhat greater than with the ethers. If the reciprocal of the time for a 25 % pressure increase is taken as a measure of the reaction velocity, the curve obtained by plotting rate against nitric oxide pressure falls to a minimum, but then slowly rises. This effect must be due to some secondary action of the nitric oxide, and it demonstrates the advantage of the initial rate over measurements made when the decomposition has proceeded to a considerable extent.

That the inhibition is a genuine effect and not due to vitiation of the pressure measurements by polymerization of the ethylene under the catalytic influence of nitric oxide is shown by two facts. The first, that when 40 mm. of nitric oxide and 100 mm. of ethylene were left for 5 min. at 660° C., no pressure decrease was observed, but, in fact, a slight increase. Secondly, the pressure change when equilibrium is reached is increased by 2 % at 660° C. for an initial ethane pressure of 150 mm. by adding 40 mm. of nitric oxide. Although there is no sudden increase in rate such as is found with dimethyl ether, the inhibitor is slowly removed in the course of the reaction. This was shown by testing samples of the gas mixture withdrawn at various stages of the reaction with a saturated solution of ferrous sulphate. Even when equilibrium has been practically attained, however, the destruction of the nitric oxide is not complete. Its removal was found by direct experiment to be due largely to the ethylene formed in the decomposition.

For 150 mm. of ethane at 620° C. the mean chain length, as given by the ratio of the rates of the uninhibited and fully inhibited reactions, is 12·2 units. The quantity of nitric oxide required to exert half the maximum inhibition is 0·62 mm. If we make the approximate assumption that at this stage a radical has an equal chance of being removed by nitric oxide or of continuing the chain, then if every collision between the inhibitor and a radical leads to the removal of the latter, only one collision in 250 between a radical and an ethane molecule is effective. This gives a value of about 10,000 cal. for the activation energy of the chain-propagating process. A value of the same order has been found for the decomposition of ethers.

The variation of the mean chain length with ethane pressure was studied at 600° C. in the unpacked reaction vessel, and also in a bulb packed with silica tubes with a surface/volume ratio 6·2 times greater. The results are given in Table II and are plotted in fig. 2. Each value of the mean chain length is taken from several determinations of the rates of the uninhibited and fully inhibited reactions. In the unpacked bulb the apparent length increases considerably as the pressure falls, but at low pressures reaches a maximum and then decreases. In the packed bulb the apparent mean chain length is everywhere less, and the maximum is displaced towards higher pressures. A direct comparison of the rate of the fully inhibited reaction in the two bulbs showed that part of the decomposition is heterogeneous. Thus, with 100 mm. of ethane, the velocity constant in sec.$^{-1} \times 10^{-4}$

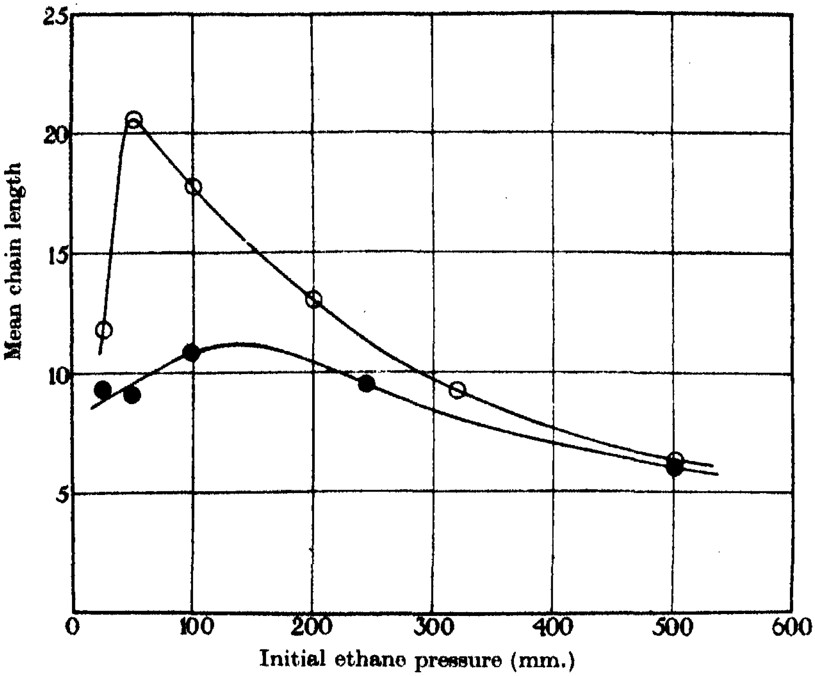


Fig. 2—Variation of the mean chain length with ethane pressure. ○ in unpacked reaction bulb; ● in tube-packed bulb.

Table II—Variation of the Mean Chain Length with the Ethane Pressure at 600° C.

In the unpacked reaction bulb		In the tube-packed bulb	
Ethane pressure (mm.)	Mean chain length	Ethane pressure (mm.)	Mean chain length
25·5	11·8	25·5	9·3
51	20·6	49·7	9·1
101·5	17·8	100·2	10·9
202	13·1	245·5	9·6
321	9·3	501·5	6·05
501·5	6·4		

in the unpacked bulb was 1·38, and in the packed vessel 1·78. At low pressures the order of both the inhibited and uninhibited reactions is considerably greater than unity. If we assume that the surface reaction depends on a lower power of the ethane concentration, and is moreover unaffected by nitric oxide, then it will form an increasing proportion of the total decomposition the lower the pressure, and more in the packed than in the unpacked bulb. In this way the difference between the two curves in fig. 2 can be satisfactorily explained. It is concluded that if there were no heterogeneous reaction the mean chain length would increase continuously as the pressure falls, and that the variation would be slightly greater than that actually observed in the unpacked reaction vessel. With dimethyl and diethyl ether the mean chain length is independent of pressure. The possible significance of this difference between ethane and the ethers will be discussed later.

For the normal decomposition of ethane Sachsse (1936*a*) found that there was an approximately linear relation between the reciprocal of the velocity constant and the reciprocal of the pressure. Such a relation is to be expected if the decomposition probability of an activated ethane molecule is independent of its energy. The discovery of the variation of the mean chain length with pressure raised the possibility that the relation found by Sachsse might be fortuitous, and it was therefore considered profitable to study the pressure dependence of the rate of the chain-free reaction. This was done at 620° C. over a pressure range of nearly 600 mm. Two experiments were made at each pressure with sufficient nitric oxide to ensure maximum inhibition. The results are given in Table III and are plotted in fig. 3. It will be seen from the figure that the order of the reaction has scarcely become unity even at 600 mm. pressure, and over a considerable range it approximates to 1·5. Also, the plot of the reciprocal of the velocity constant against that of the pressure, far from being linear, bends over rapidly towards the pressure axis. The curvature of the line would certainly be less if it were possible to make allowance for the surface reaction, but it is doubtful whether the linear relation found by Sachsse holds for the chain-free decomposition.

The temperature coefficient of the inhibited reaction was found by measurements of the initial rate at six temperatures varying over a range of 100° C. At least two determinations were made at each temperature, care being taken to add sufficient nitric oxide to produce the maximum possible inhibition. 500 mm. of ethane were used throughout. At this pressure the same value for the mean chain length was found in the unpacked bulb, the tube-packed bulb previously referred to, and in a ballpacked bulb with a surface/volume ratio 11·5 times that of the unpacked vessel. The surface reaction can therefore only make a very small contribution to the total rate of the decomposition. The experimental results are given in Table IV, where they are compared with the values calculated from the equation

$$\log_e k = 33{\cdot}21 - 74{,}500/RT.$$

The energy of activation, 74,500 cal., seems to be the largest recorded for

TABLE III—VARIATION OF THE RATE OF THE INHIBITED REACTION WITH INITIAL ETHANE PRESSURE

Ethane pressure (mm.)	Velocity constant (sec.$^{-1}$ × 10^6)
10·7	3·08
20·7	4·26
43·2	5·20
80·7	6·78
141·2	9·05
199·5	11·2
306	15·5
400	16·8
501	17·7
596	19·4

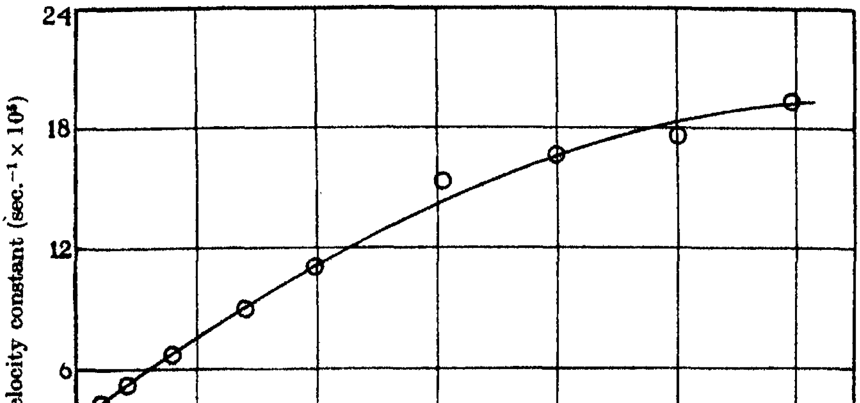


FIG. 3—Variation of the initial rate of the inhibited reaction with the ethane pressure.

the decomposition of an organic compound. At 620° C. and with an initial ethane pressure of 500 mm. the observed velocity constant of the chain-free reaction is $1{\cdot}235 \times 10^{-4}$ sec.$^{-1}$, which gives $6{\cdot}67 \times 10^{14}$ for the number of molecules reacting per c.c. per sec. Taking the diameter of the ethane molecule as 4×10^{-8} cm. the value of $Ze^{-E/RT}$ is $9{\cdot}0 \times 10^{9}$ (Z being the number of collisions per c.c. per sec.). The number of molecules actually reacting is therefore $7{\cdot}4 \times 10^{4}$ times greater than this. Using the expression of Hinshelwood (1926) the number of square terms in which the activation energy must be assumed to be distributed in order to account for this factor is 14. This is perhaps large for such a relatively simple molecule, but not unreasonably so.

TABLE IV—TEMPERATURE COEFFICIENT OF THE INHIBITED REACTION

Temperature (° C.)	$\log_{10}(k \times 10^{6})$	
	Observed	Calculated
560	0·88	0·80
580	1·28	1·27
600	1·69	1·70
620	2·09	2·13
640	2·53	2·53
660	2·94	2·91

It was found that ethylene can appreciably reduce the initial rate of the decomposition, though the effect is not of the same order as that shown by nitric oxide. For an initial ethane pressure of 150 mm. at 660° C., the velocity constant of the normal decomposition is $3{\cdot}66 \times 10^{-3}$ sec.$^{-1}$. This is reduced to $2{\cdot}02 \times 10^{-3}$ by the addition of 68 mm. of ethylene, and to $1{\cdot}43 \times 10^{-3}$ by 101 mm. Unpublished observations by Mitchell have shown that the radicals produced in the photolysis of acetaldehyde bring about the rapid polymerization of ethylene. Such a polymerization under the influence of the radicals liberated from the ethane may account for part of the effect.

DISCUSSION

Where the nitric-oxide test has revealed the presence of reaction chains in the thermal decompositions of organic compounds it has been concluded that only a small proportion of the activated molecules yield radicals, the great majority undergoing rearrangement processes. The chief reason for thinking this is that in all these cases the activation energy of the chain-free reaction is considerably less than the energy required to break the carbon-carbon or carbon-oxygen bond. For ethane the activation energy of the chain-free reaction is 74,500 cal., which is considerably higher than that for, say, any of the ethers. However, even though the carbon-carbon bond energy may be only 10,000–20,000 cal. greater than this, it does not follow that a relatively large fraction of activated ethane molecules decompose into radicals, since the activation energy is known to be widely distributed in about 14 square terms, and the chance that it all migrates into the carbon-carbon link before the molecule is deactivated may be very small indeed. It seems useless to attempt any quantitative calculations of the fraction of activated molecules which yield radicals until more is known about the distribution of energy in molecules and of the exact values of bond strengths. By determining the rate at which the conversion of para to ortho hydrogen takes place in the presence of decomposing ethane, Sachsse (1936*b*) has concluded that only one in every 10^4 to 10^6 activated ethane molecules gives radicals. With a mean chain length of the order of 10 units, the absolute chain length would be from 10^5 to 10^7 units. At least we can say that the evidence strongly indicates that the chains are few in number and long.

According to Rice and Herzfeld (1934), the methyl radicals produced when the carbon-carbon bond in ethane is broken react with ethane to give methane and an ethyl radical. This decomposes into an ethylene molecule and a hydrogen atom, which brings about the further decomposition of ethane by the following chain cycle:

$$H + C_2H_6 = C_2H_5 + H_2,$$

$$C_2H_5 = C_2H_4 + H.$$

If this mechanism is correct, then hydrogen atoms are the chain carriers, and their removal from the system may either take place on the walls of the reaction vessel or in the gas phase. The absence of marked wall effects in the uninhibited decomposition supports the second alternative. It is interesting to enquire how the mean chain length will vary with the ethane pressure according to whether one assumes that in the absence of nitric oxide the chain carriers are removed from the system by binary or by ternary collisions. Let a be the concentration of the activated molecules which yield radicals, n that of the ethane molecules, and R that of the radical or atom which propagates the chain. We may assume for simplicity that in the propagation of the chain radicals are regenerated as fast as they are removed. If the chains are terminated by the recombination of two chain-propagating particles in the presence of a third body (an ethane molecule), then for a stationary radical concentration we have

$$dR/dt = k_1 a - k_2 R^2 n = 0,$$

whence

$$R = \sqrt{(k_1 a/k_2 n)},$$

where k_1 and k_2 are constants. a will depend on some function $f_1(n)$ of the ethane concentration, and can, in fact, be put proportional to a power of n between 1 and 2. The rate of the rearrangement processes will also depend on some function of n, say $f_2(n)$. Hence the rate of the uninhibited reaction will be given by the expression

$$k_3 f_2(n) + k_4 Rn$$
$$= k_3 f_2(n) + k_4 \sqrt{(k_1 n . f_1(n)/k_2)},$$

and that of the inhibited reaction simply by $k_3 f_2(n)$, where k_3 and k_4 are constants. The mean chain length, being the ratio of these two rates, will be

$$1 + K\sqrt{(n . f_1(n)/[f_2(n)]^2)},$$

where K is a constant. If, however, the chain carriers recombine by binary collisions in the gas phase, this expression for the mean chain length becomes

$$1 + K\sqrt{(n^2 . f_1(n)/[f_2(n)]^2)}.$$

Over a considerable range of pressure the order of the inhibited reaction is approximately 1·5, so that we may put $f_2(n)$ equal to $n^{1\cdot 5}$. Since $f_1(n)$ can be represented by n^x, where x is between 1 and 2, we see that if the chains end by binary recombination of radicals the mean chain length must either increase with increasing pressure or remain constant, whereas on the alternative assumption of ternary collisions a decrease with increasing pressure is possible. The chance that a molecule with sufficient energy dissociates into radicals before deactivation may well be very small indeed, as the energy must become localized in one bond. According to accepted views on the mechanism of activation in unimolecular reactions, when the decomposition probability of the activated molecule is very small, the decomposition process is of the first order except at extremely low pressures. It is not unlikely, therefore, that in the pressure region where the order of the rearrangement processes is about 1·5 the value of x is less than this, and perhaps almost unity. In the latter event the mean chain length if ternary collisions are all-important would vary inversely as the square root of the ethane pressure. The observed dependence is roughly of this order.

If methyl or ethyl radicals somehow propagate the chains, there is no reason why they should not recombine simply by binary collisions in the gas phase, and there is in fact evidence that when they are the chain carriers they are removed in this way. (See *inter alia* Mitchell and Hinshelwood 1937; Allen and Sickman 1934; Fletcher and Rollefson 1936.) The presence of a third body, however, is essential for the recombination of hydrogen atoms. It is thus concluded tentatively that in this reaction hydrogen atoms are the chain carriers. It is of course possible that the chains are terminated by the recombination of hydrogen atoms and ethyl radicals. If it is assumed that the decomposition of ethyl radicals in the chain cycle into hydrogen atoms and ethylene results from their collision with ethane molecules, then a detailed analysis of the problem based on the scheme of Rice and Herzfeld (1934) shows that the decrease in the mean chain length with increasing pressure is again best explained if the presence of a third body is essential for the recombination. On general grounds a hydrogen atom and an ethyl radical are less likely to unite in a simple binary process than, say, two ethyl radicals.

At least there seems to be a significant difference between ethane and dimethyl and diethyl ethers. With ethane there is a pronounced decrease in the mean chain length with increasing pressure. Hydrogen atoms seem to be the most likely chain carriers. With the ethers the mean chain length is independent of pressure, and probably methyl or ethyl radicals propagate the chains.

I am indebted to Professor C. N. Hinshelwood for his much valued advice in connexion with this work.

SUMMARY

Nitric oxide reduces the initial rate of the thermal decomposition of ethane to a well-defined limit. This effect is attributed to the ability of nitric oxide to suppress reaction chains by combining with free radicals or atoms. The values found for the mean chain length (the ratio of the rates of the uninhibited and fully inhibited reactions) vary between 6 and 21, according to the temperature and pressure. It seems probable that only a small proportion of activated ethane molecules yield radicals, the great majority undergoing direct rearrangement to give ethylene and hydrogen. The absolute chain length must therefore be high.

The mean chain length decreases considerably as the initial ethane pressure increases. This fact is discussed in the light of the possibility that hydrogen atoms are the chain carriers. The activation energy of the chain propagation process is about 10,000 cal.

The activation energy of the chain-free reaction has the very high value of 74,500 cal., and is distributed in about 14 square terms.

References

Allen and Sickman 1934 *J. Amer. Chem. Soc.* **56**, 2031.
Fletcher and Rollefson 1936 *J. Amer. Chem. Soc.* **58**, 2135.
Hinshelwood 1926 *Proc. Roy. Soc.* A, **113**, 230.
Mitchell and Hinshelwood 1937 *Proc. Roy. Soc.* A, **159**, 32.
Rice and Herzfeld 1934 *J. Amer. Chem. Soc.* **56**, 284.
Sachsse 1936*a* *Z. phys. Chem.* B, **31**, 87.
— 1936*b* *Z. phys. Chem.* B, **31**, 79.
Staveley and Hinshelwood 1936*a* *Nature, Lond.*, **137**, 29.
— 1936*b* *Proc. Roy. Soc.* A, **154**, 335.
— 1936*c* *J. Chem. Soc.* pp. 812, 818.
— 1937 *Proc. Roy. Soc.* A, **159**, 192.

X-Ray Analysis of the Dibenzyl Series
IV—Detailed Structure of Stilbene

By J. Monteath Robertson, M.A., D.Sc., and I. Woodward, M.A.

(*Communicated by Sir William Bragg, O.M., P.R.S.—Received* 10 *July* 1937)

A preliminary account of the crystal structure of stilbene in relation to other compounds in the dibenzyl series has already been given in Part III of this work (Robertson, Prasad and Woodward 1936). The general arrangement of the four molecules in the unit cell was deduced from certain absolute intensity measurements, the physical properties of the crystal, and a comparison with the accurately known structure of dibenzyl (Robertson 1934, 1935*a*). Although the approximate structure was established beyond any doubt, the precise determination of the atomic positions has proved to be a matter of very great difficulty. This is due to the presence of four molecules in the unit cell and to the fact that it is impossible to formulate a definite model of the molecule from the chemical structure owing to the possibility of free rotation about the single bonds. For the accurate analysis it was necessary to determine the phase constants of all the weaker réflexions, and in the course of this work over twenty trial structures were calculated and compared. In this connexion the mechanical sorting method of structure factor calculation previously described (Robertson 1936) proved of great assistance. The repeated application of a double Fourier synthesis to the $(h0l)$ set of structure factors then gradually refined the values for the atomic co-ordinates, and we believe that the final results given below are now comparable in accuracy to those obtained for the other structures in this series. It is unnecessary to describe in detail the various stages by which this structure has been reached, and the results of the final Fourier synthesis only are given in the following pages.

The relative orientations of the molecules in the crystal now recorded should enable various physical properties of the crystal to be expressed as functions of the individual molecules. In particular, the results now obtained will justify more accurate measurements being made on the magnetic susceptibilities and the refractive indices of the crystal than are available at present.

The main results of the analysis are first discussed in relation to the structure of the molecule, while the X-ray work, co-ordinates of the atoms, etc., are given in a later section.

THE STRUCTURE OF THE STILBENE MOLECULE

Two separate molecules contribute to the asymmetric unit of the crystal, and their structures have been independently deduced from the intensity measurements. The final results are expressed by contour maps of the electron distribution in figs. 1, 2 and 3 which represent projections of the structure along the b crystal axis. It will be seen that the two molecules differ considerably in these projections. It is found, however, that this difference is due to relative orientation and not to any significant difference in the shape or dimensions of the molecules themselves. The actual dimensions can readily be deduced from these projections, and the results are shown in figs. 4 and 5.

Owing to the complicated nature of the structure only the b axis projections give any resolution of separate atoms. Consequently as only one projection is available for each molecule, some of the dimensions recorded in figs. 4 and 5 are only partially direct determinations. This applies particularly to the carbon-carbon double bond distance of 1·33 A and to

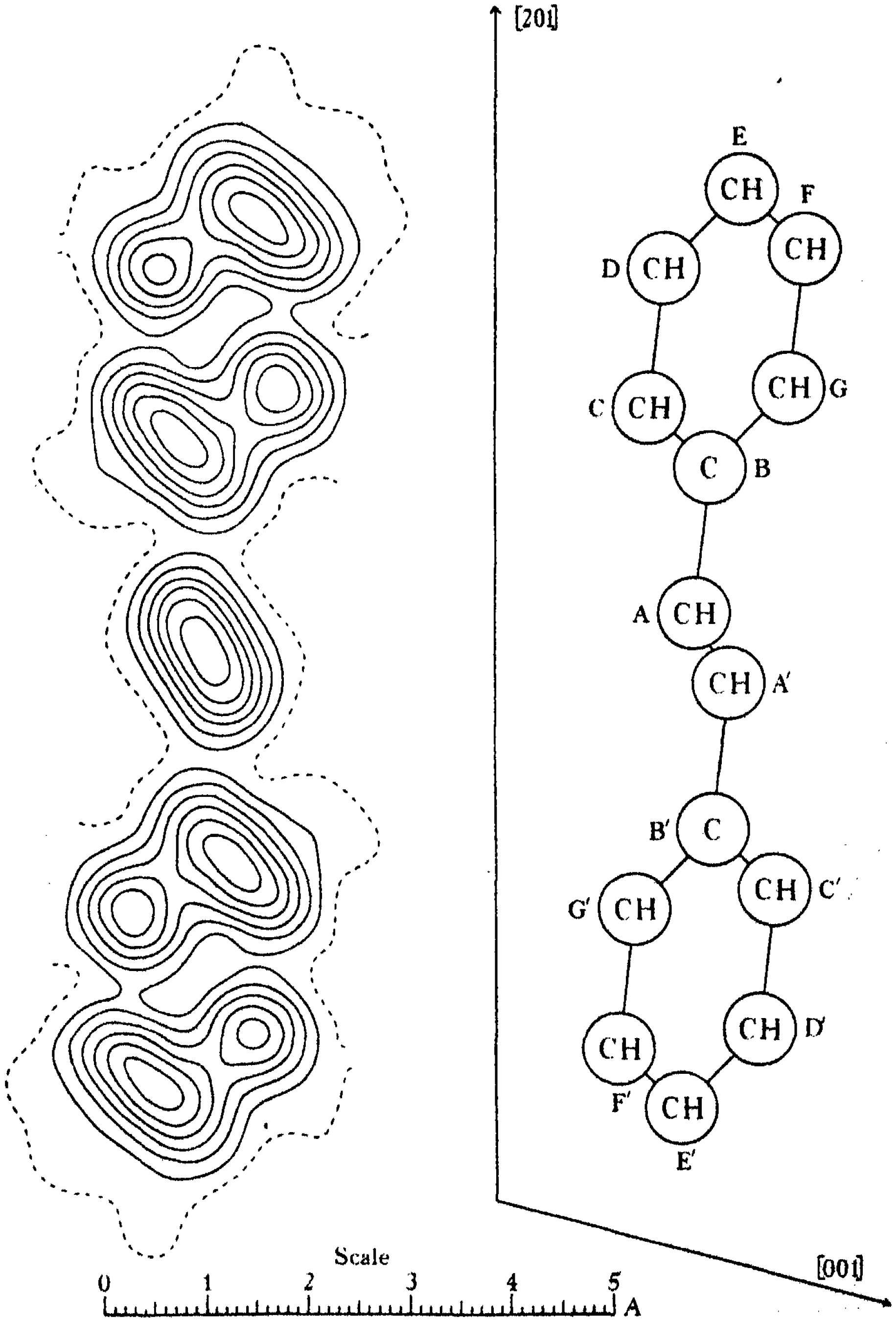


FIG. 1—Stilbene molecule 1 in the *b* axis projection. Each contour line represents a density increment of approximately one electron per A^3, the one electron line being dotted.

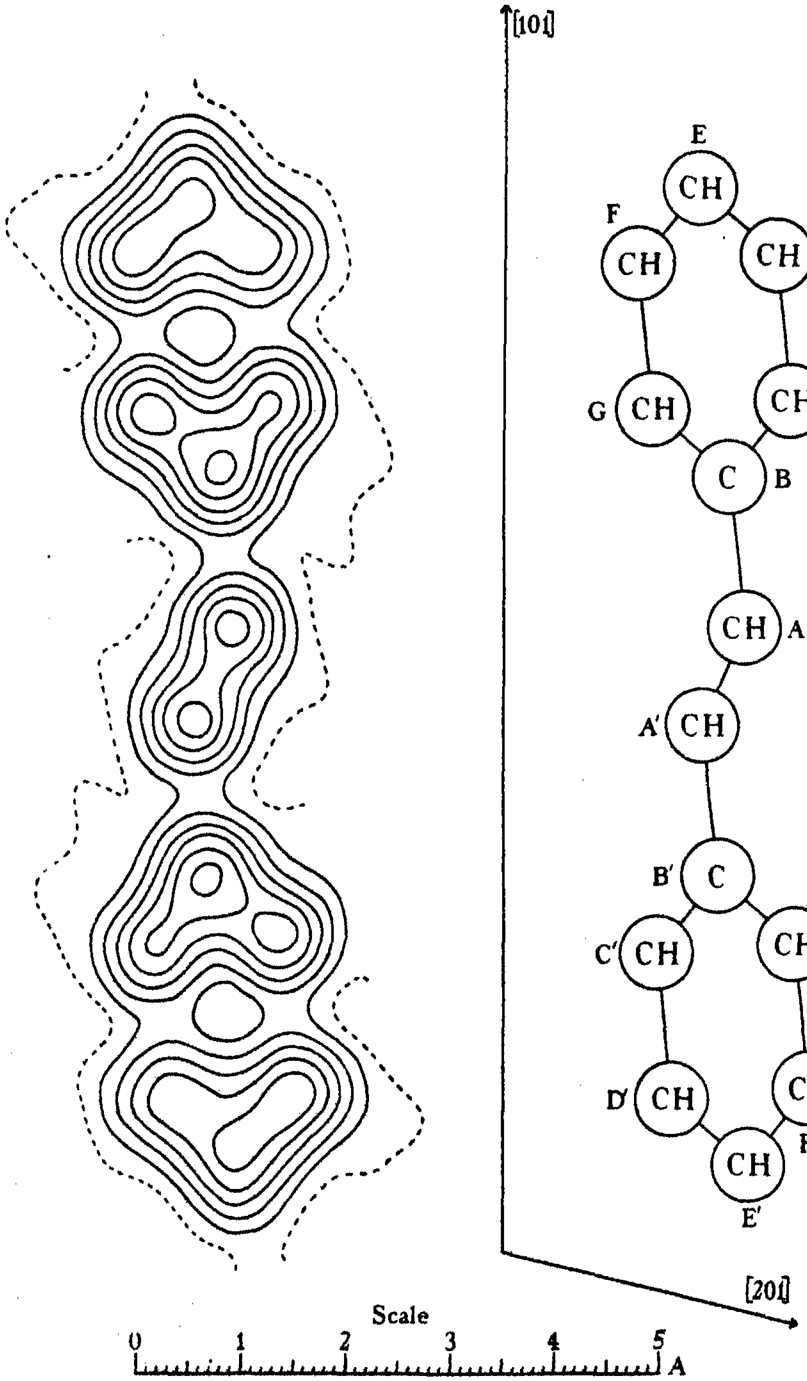


Fig. 2—Stilbene molecule 2 in the *b* axis projection. Scale as in fig. 1.

the carbon-carbon distance in the ring of 1·39 A. These values were assumed in calculating the orientations of the molecules, and were later justified by the agreements obtained when other sets of structure factors were calculated. The double bond distance in particular is subject to a possible error of about ± 0·04 A. The "single bond" distance between the benzene ring and the carbon atom *A* is a much more direct determination owing to the small inclination which this line makes with the plane of the projection. For this distance we obtain the value 1·44 A in one molecule and 1·45 A

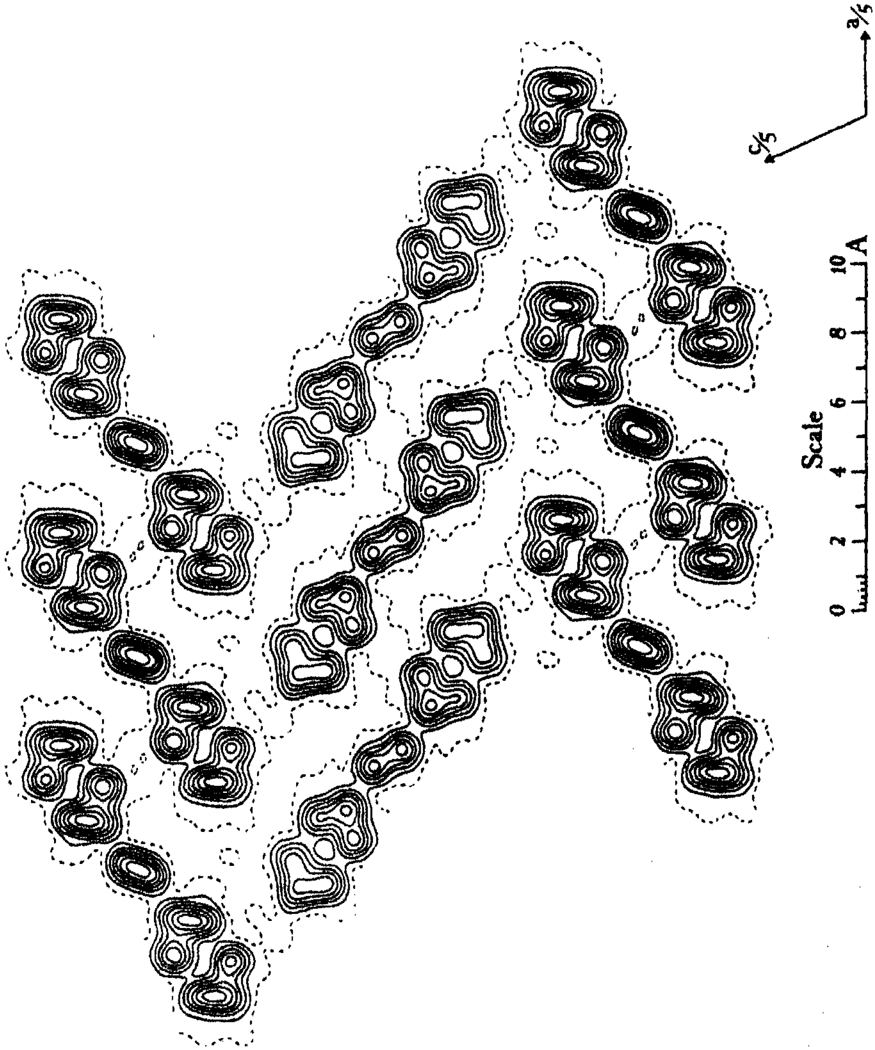
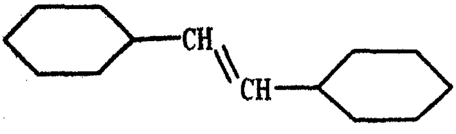


FIG. 3—Showing on a smaller scale the arrangement of a group of stilbene molecules in the *b* axis projection. Each line approximately one electron per A^2.

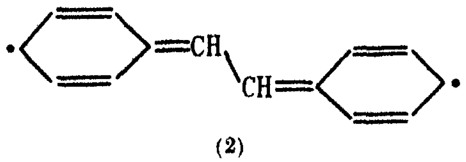
in the other, figures which, however, are still subject to an error of at least ± 0·02 A, so that we cannot attach any significance to the difference.

With regard to shape and the angles between the bonds, a somewhat larger difference is obtained between the two molecules. They are both nearly flat. The divergence from a strictly planar model can be measured by the angle through which it is necessary to rotate the ring about the single bond AB, starting from the planar position. For molecule 1 this rotation is 10° and for molecule 2 only 3°. Further, the angle between the double bond and the single bond ($A'AB$) is 128° for molecule 1 and 133° for molecule 2. It is doubtful, however, whether much significance can be attached to these observed differences. The short distance between the central atoms A and A' and their imperfect resolution in the projections make it difficult to define the exact direction of the line joining them, and all the above figures depend upon the orientation of this line. It seems possible that strictly planar models may apply to each molecule, within the limits of experimental error.

In spite of these minor uncertainties definite conclusions can be drawn regarding the structure of the stilbene molecule. The length of the single bond from the benzene ring to the atom A has undergone a severe contraction from the standard value of 1·54 A as in diamond. The value of 1·44 A which we find indicates about 40 % of double bond character in this link, if use is made of the curve recently given by Penney (1937) which expresses the relation between the order of linkage and the internuclear distance between two linked carbon atoms (the "ionic diameter" of carbon). There is thus no doubt that the ordinary formula for stilbene (1) does not represent the molecule completely, but that excited structures of the form represented by (2) must also contribute to the normal state. In (2) the dots in the para positions represent single electrons, which can formally be considered as paired. Such formulae have been discussed by Pauling and Sherman (1933) in relation to diphenyl and related compounds.



(1)



(2)

The strongest evidence for the existence of (2), however, is to be found in the planar form of the molecule. In (1) free rotation of the rings is possible, and in the nearly isomorphous dibenzyl crystal the planes of the benzene rings are in fact practically at right angles to the plane of the central zig-zag. (The rings are always parallel to each other, both in the dibenzyl and stilbene crystals, owing to the centre of symmetry in the molecule.) Thus as far as formula (1) is concerned we might expect the plane of the rings to be at right angles to the plane of the central zig-zag, by analogy with the dibenzyl crystal, and especially as such a configuration would give a greater clearance between the ethylene hydrogen atoms and the ortho hydrogen atoms in the ring. But if formula (2) contributes to the structure, the molecule is necessarily planar, from general stereochemical considerations.

We must conclude that in the actual structure there is resonance between these two forms, and their relative contributions may be estimated from the internuclear distances given above. Further evidence for the existence of such resonance in stilbene has been given by Pauling and Sherman (1933) from a study of thermochemical data. They find that the total resonance energy of stilbene is 4·09 e-volts, which is 0·67 e-volt more than the resonance energy for two benzene rings. This additional energy must be attributed to resonance between (1) and (2).

With regard to the unusually large angle of 128–133° which we find in the central zig-zag of stilbene (compare figs. 4 and 5), we are indebted to Dr. Jahn for pointing out that this is probably a result of repulsion between the ortho hydrogen atoms in the ring and the ethylene hydrogen atoms. If a planar model is made, with the hydrogen atoms at 1·09 A from the carbon atoms and all the angles 120° (as in ethylene), then the distance between the ortho hydrogen atom and the ethylene hydrogen atom is 1·71 A on one side of the ring and 2·45 A on the other side. Increasing the zig-zag angle to 133° makes these distances 2·33 A and 2·07 A respectively. As the carbon-hydrogen bonds themselves must also react to the repulsive forces to some extent, it is probable that with a zig-zag angle of about 130° the hydrogen-hydrogen distances are almost equalized. The increase in the angle might, however, also arise from a repulsion between the π electrons on non-neighbouring carbon atoms (atoms A and B' or A' and B). This question could be settled from measurements on compounds with longer conjugated chains, such as the diphenylpolyenes.

ANALYSIS OF THE STRUCTURE

Crystal Data

Stilbene—$C_{14}H_{12}$; melting point 124° C.; density, measured 1·159 (Hengstenberg), calculated 1·161; monoclinic prismatic, $a = 12{\cdot}35$, $b = 5{\cdot}70$, $c = 15{\cdot}92$ A, $\beta = 114{\cdot}0°$. Space group $C_{2h}^5(P2_1/a)$. Four molecules per unit cell. Molecular symmetry, centre. Volume of the unit cell, 1023 A^3. Absorption coefficient for X-rays, $\lambda = 1{\cdot}54$, $\mu = 5{\cdot}97$ per cm. Total number of electrons per unit cell $= F(000) = 384$.

Fourier Synthesis

The X-ray measurements on stilbene have been described in a previous paper (Robertson, Prasad and Woodward 1936). The absolute scale of the F values is difficult to determine accurately for crystals in this series of compounds owing to their volatile nature. This is particularly the case for azobenzene, the unprotected crystal rapidly losing weight at room temperatures. On the other hand, the use of sealed containers for the crystals tends to reduce the accuracy of the intensity measurements. Because of this the scale of the F values was finally decided by making a correlation with the final calculated values, which are based on an average atomic f curve for aromatic hydrocarbons (Robertson 1935*b*). Direct measurements on stilbene gave results which tended to be rather lower than this scale.

In the analysis of the structure one of the chief difficulties lay in determining the direction of the central —CH=CH— groups in the two molecules which contribute to the asymmetric unit. After a number of trial calculations, a preliminary Fourier synthesis was carried out, in which all terms whose sign depended upon the directions assigned to these groups were omitted. The result gave a rather ill-defined projection of the structure but indicated quite clearly the general direction of the central groups and the lie of the rings. It was then possible to calculate the phase constants of further terms in the series and gradually refine the structure.

The coefficients in the double Fourier series used for the contour maps of figs. 1–3 are the measured values of $F(h0l)$ given in Table I, divided by the area of the (010) face, with signs corresponding to the calculated values. From a total of a hundred possible terms six of the smaller ones, amounting to about 2 % of the total, were omitted from the final synthesis because of doubtful sign. The series was summed at 900 different points on the asymmetric unit (one-quarter of the unit cell); at intervals of 6° or 0·206 A on the a axis, and 6° or 0·265 A on the c axis. The numerical results are not reproduced in this paper, but the contour maps were drawn by graphical interpolation methods from the summation totals.

Orientation of the Molecules. Co-ordinates

In both molecules the benzene rings correspond fairly accurately to the projections of regular plane hexagons. In molecule I (fig. 1) two of the atoms, D and G, are quite clearly defined, but the atoms of the central

—CH=CH— group are not resolved. In molecule 2 (fig. 2) the atoms of the ring are more obscure, but the central group is partially resolved. The orientation of the molecules can best be calculated by measuring the angles which certain well-defined lines in the projections make with the crystal axes, and assuming certain known interatomic distances and valency angles.

TABLE I—VALUES OF $F(h0l)$

				h			
	0	2	4	6	8	10	12
12	—	—	−10	—	—	—	—
11	—	+7	—	—	—	—	—
10	+8	−14	—	—	—	—	—
9	—	+14	−7	—	—	—	—
8	−20	+12	—	—	—	—	—
7	+6	+14	−18	+9	−6	—	—
6	+13	+19	−9	—	—	—	—
5	—	—	—	—	−15	—	—
4	−57	—	+40	+5	+9	—	—
3	—	−53	+33	−30	—	—	—
2	+24	+89	+37	+7	—	−4	—
1	—	−53	—	−4	+9	—	—
0	+384	−37	—	+14	+6	—	—
$\bar{1}$	—	—	−17	+32	—	—	−8
$\bar{2}$	+24	−36	+18	+18	+17	−6	+11
$\bar{3}$	—	+50	+20	—	—	−12	—
$\bar{4}$	−57	+81	—	+27	+8	+17	—
$\bar{5}$	—	+53	+7	−38	—	+11	—
$\bar{6}$	+13	+8	+39	+14	—	+20	−18
$\bar{7}$	+6	+6	−26	+15	+8	+12	−8
$\bar{8}$	−20	+12	+31	—	—	—	—
$\bar{9}$	—	—	—	+35	−10	—	+14
$\bar{10}$	+8	+8	−5	+9	—	—	+10
$\bar{11}$	—	−5	+11	—	—	+8	+11
$\bar{12}$	—	−9	+8	−5	+9	−6	—
$\bar{13}$	—	−12	+14	—	+19	—	—
$\bar{14}$	—	—	−9	—	—	—	—
$\bar{15}$	—	+6	+6	—	+5	—	—
$\bar{16}$	—	−8	−7	—	—	—	—

The quantities used are defined as follows:

r = measured length of any line in the projection.

R = assumed real length.

η = angle which projected line makes with a crystal axis.

(χ, ψ, ω) = the angles which the line actually makes with the a, b and c' crystal axes, c' being perpendicular to a and b.

L, M, and N, are three convenient axes in the molecule, mutually perpendicular (compare fig. 4).

Then we have the relations

$$(1)\ \cos^2\chi+\cos^2\psi+\cos^2\omega = 1,$$

$$(2)\ \sin\psi = r/R,$$

$$(3)\ \cos\omega = \cos\chi\tan\eta.$$

For molecule 1 (compare figs. 1 and 6) the measured quantities are

$$r_{DG} = 1{\cdot}68,\quad r_{AA'} = 0{\cdot}76,\quad r_{AB} = 1{\cdot}43\text{ A},$$

$$\eta_{DG} = 84{\cdot}8^\circ,\quad \eta_{AA'} = 65{\cdot}6^\circ,\quad \eta_{L(AB)} = 32{\cdot}0^\circ.$$

We further assume that $R_{DG} = 2{\cdot}78$ A, $R_{AA'} = 1{\cdot}33$ A (double bond distance in ethylene) and that the angle between L and DG in the benzene ring is 60°. The axis M is then taken to lie in the plane of L and DG.

These figures give the following values for the orientation of molecule 1:

$\chi_L = 33{\cdot}4^\circ$ $\cos\chi_L = 0{\cdot}8352$ $\chi_M = 114{\cdot}8^\circ$ $\cos\chi_M = -0{\cdot}4190$
$\psi_L = 79{\cdot}9^\circ$ $\cos\psi_L = 0{\cdot}1748$ $\psi_M = 35{\cdot}0^\circ$ $\cos\psi_M = 0{\cdot}8187$
$\omega_L = 58{\cdot}5^\circ$ $\cos\omega_L = 0{\cdot}5219$ $\omega_M = 66{\cdot}8^\circ$ $\cos\omega_M = 0{\cdot}3939$

$\chi_N = 110{\cdot}9^\circ$ $\cos\chi_N = -0{\cdot}3563$ $\chi_{DG} = 86{\cdot}9^\circ$ $\cos\chi_{DG} = 0{\cdot}0547$
$\psi_N = 123{\cdot}2^\circ$ $\cos\psi_N = -0{\cdot}5469$ $\psi_{DG} = 37{\cdot}2^\circ$ $\cos\psi_{DG} = 0{\cdot}7965$
$\omega_N = 40{\cdot}8^\circ$ $\cos\omega_N = 0{\cdot}7567$ $\omega_{DG} = 53{\cdot}0^\circ$ $\cos\omega_{DG} = 0{\cdot}6021$

$\chi_{AA'} = 76{\cdot}3^\circ$ $\cos\chi_{AA'} = 0{\cdot}2362$
$\psi_{AA'} = 34{\cdot}9^\circ$ $\cos\psi_{AA'} = 0{\cdot}8204$
$\omega_{AA'} = 58{\cdot}6^\circ$ $\cos\omega_{AA'} = 0{\cdot}5207$

The angles which the central —CH═CH— group (AA') makes with the L, M and N axes are 127·8°, 39°, and 98°. Thus the angle between AA' and the plane of the ring is 8·0°, and the angle through which the ring must be turned about AB or L, starting from a flat model, is 10·4°. The reduced dimensions of the molecule are shown in fig. 4.

The rectangular co-ordinates of the atom A, referred to the a, b and c' crystal axes and centre of symmetry (mid point of AA') as origin, are

$$x'_A = 0{\cdot}157,\quad y'_A = 0{\cdot}546,\quad z'_A = 0{\cdot}346\text{ A},$$

while the co-ordinates of the other atoms, referred to the molecular axes L and M with A as origin, are given in Table II.

TABLE II—MOLECULE 1

Atom ...	A	B	C	D	E	F	G
L ...	0	$1{\cdot}45_1$	$2{\cdot}14_6$	$3{\cdot}53_6$	$4{\cdot}23_1$	$3{\cdot}53_6$	$2{\cdot}14_6$
M ...	0	0.	$1{\cdot}20_3$	$1{\cdot}20_3$	0	$-1{\cdot}20_3$	$-1{\cdot}20_3$

It has already been shown (Robertson, Prasad and Woodward 1936) that molecule 2 can be derived from molecule 1 approximately by a rotation of 180° about the *a* axis and a translation of *c*/2. Hence in order to compare the relative orientations of the two molecules in the crystal the above movement is applied to molecule 2 before calculating its orientation. With this provision, the calculations are very similar to the above. The line *DG*, however, is not so clearly defined in molecule 2, so the projection of a regular hexagon was first described on the ring to give the best fit at all the corners, and the subsidiary line, *PQ*, parallel to *M*, was used.

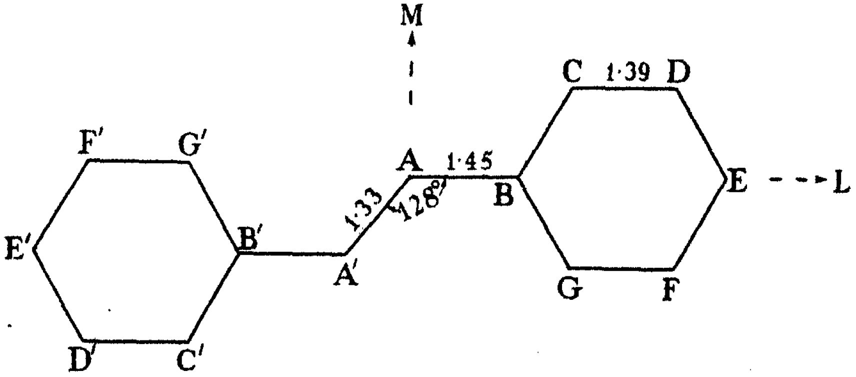


Fig. 4—Dimensions of stilbene molecule 1. *AA′* is inclined at 8° to plane of ring.

The measured quantities for molecule 2 are then

$$r_{PQ} = 1{\cdot}33, \qquad r_{AA'} = 1{\cdot}02, \qquad r_{AB} = 1{\cdot}44 \text{ A},$$

$$\eta_{M(PQ)} = 123{\cdot}8^\circ, \quad \eta_{AA'} = 61{\cdot}2^\circ, \quad \eta_{L(AB)} = 31{\cdot}4^\circ.$$

It is further assumed that $R_{PQ} = 2 \times 1{\cdot}39 \ \cos 30^\circ = 2{\cdot}406$ A, and that $AA' = 1{\cdot}33$ A as before.

With these figures the following values are obtained for the orientation of molecule 2:

$\chi_L = 31{\cdot}4^\circ$ $\quad \cos\chi_L = 0{\cdot}8533$ $\qquad \chi_M = 107{\cdot}9^\circ$ $\quad \cos\chi_M = -0{\cdot}3076$

$\psi_L = 88{\cdot}4^\circ$ $\quad \cos\psi_L = 0{\cdot}0278$ $\qquad \psi_M = 33{\cdot}6^\circ$ $\quad \cos\psi_M = 0{\cdot}8333$

$\omega_L = 58{\cdot}6^\circ$ $\quad \cos\omega_L = 0{\cdot}5208$ $\qquad \omega_M = 62{\cdot}6^\circ$ $\quad \cos\omega_M = 0{\cdot}4595$

$\chi_N = 114{\cdot}9^\circ$ $\quad \cos\chi_N = -0{\cdot}4212$ $\qquad \chi_{AA'} = 68{\cdot}3^\circ$ $\quad \cos\chi_{AA'} = 0{\cdot}3694$

$\psi_N = 123{\cdot}5^\circ$ $\quad \cos\psi_N = -0{\cdot}5521$ $\qquad \psi_{AA'} = 50{\cdot}1^\circ$ $\quad \cos\psi_{AA'} = 0{\cdot}6417$

$\omega_N = 44{\cdot}0^\circ$ $\quad \cos\omega_N = 0{\cdot}7194$ $\qquad \omega_{AA'} = 47{\cdot}8^\circ$ $\quad \cos\omega_{AA'} = 0{\cdot}6719$

The angles which LA' makes with L, M and N for molecule 2 are thus 133·1°, 43·1° and 91·5°, the angle between AA' and the plane of the ring, 1·5°, and the angle through which the ring must be turned about L from the strictly planar model is 2·8°. Hence molecule 2 is almost exactly flat. The reduced dimensions are shown in fig. 5.

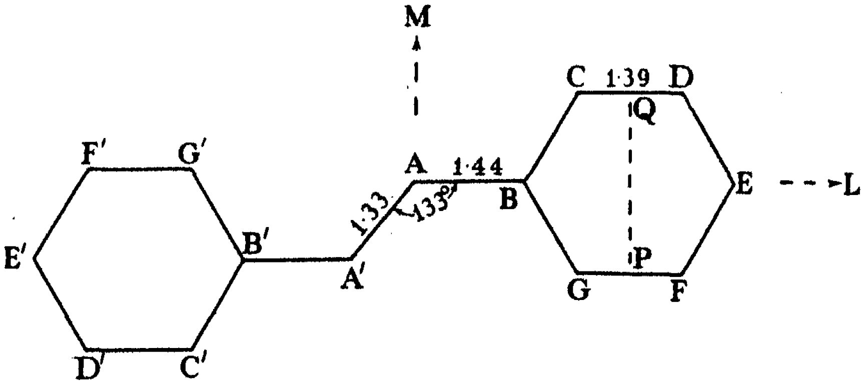


FIG. 5—Dimensions of stilbene molecule 2. AA' is inclined at 1·5° to plane of ring.

The rectangular co-ordinates of the atom A, referred to the a, b and c' crystal axes and the centre of symmetry of molecule 2 as origin, are

$$x'_A = 0{\cdot}246, \quad y'_A = 0{\cdot}427, \quad z'_A = 0{\cdot}447 \text{ A},$$

while the co-ordinates of the other atoms, referred to the molecular axes L and M with A as origin, are given in Table III.

TABLE III—MOLECULE 2

Atom...	A	B	C	D	E	F	G
L ...	0	$1{\cdot}44_0$	$2{\cdot}13_5$	$3{\cdot}52_5$	$4{\cdot}22_0$	$3{\cdot}52_5$	$2{\cdot}13_5$
M ...	0	0	$1{\cdot}20_8$	$1{\cdot}20_8$	0	$-1{\cdot}20_8$	$-1{\cdot}20_8$

The crystal co-ordinates of all the atoms in the asymmetric unit, which includes both molecules, are given in Table IV. These figures are referred to the monoclinic crystal axes and the centre of symmetry at (000) as origin. For molecule 1 they are derived from the preceding figures by the relations

$$x = x' - z' \cot \beta,$$

$$y = y',$$

$$z = z' \operatorname{cosec} \beta,$$

and for molecule 2, referred to the same origin, by

$$x = x' + z' \cot \beta,$$

$$y = -y',$$

$$z = c/2 - z' \operatorname{cosec} \beta,$$

where in both cases

$$x' = L \cos \chi_L + M \cos \chi_M + x'_A,$$

$$y' = L \cos \psi_L + M \cos \psi_M + y'_A,$$

$$z' = L \cos \omega_L + M \cos \omega_M + z'_A.$$

TABLE IV—CO-ORDINATES. CENTRE OF SYMMETRY AT (000) AS ORIGIN. MONOCLINIC CRYSTAL AXES

Atom	x	$2\pi x/a$	y	$2\pi y/b$	z	$2\pi z/c$
A_1 (CH)	0·31 A	9·1°	0·55 A	34·5°	0·38 A	8·6°
B_1 (C)	1·86	54·2	0·80	50·5	1·21	27·3
C_1 (CH)	2·31	67·3	1·91	120·3	2·12	48·0
D_1 (CH)	3·79	110·6	2·15	135·7	2·92	65·9
E_1 (CH)	4·83	140·8	1·28	81·2	2·79	63·2
F_1 (CH)	4·38	127·7	0·18	11·3	1·88	42·5
G_1 (CH)	2·89	84·4	−0·06	− 4·0	1·08	24·5
A_2 (CH)	0·05	1·4	−0·43	−27·0	7·47	169·0
B_2 (C)	0·94	27·5	−0·47	−29·5	6·65	150·3
C_2 (CH)	0·76	22·1	−1·49	−94·0	5·65	127·7
D_2 (CH)	1·62	47·2	−1·53	−96·4	4·86	109·8
E_2 (CH)	2·67	77·8	−0·54	−34·4	5·07	114·6
F_2 (CH)	2·85	83·2	0·48	30·1	6·07	137·2
G_2 (CH)	1·99	58·0	0·52	32·6	6·86	155·1

The manner in which the co-ordinates finally assigned to the atoms fit the contour maps is shown by the plot in fig. 6, which covers the asymmetric portion of the unit cell.

Intermolecular Distances and Arrangement of Molecules

The minimum distance of approach between atoms on adjoining molecules in the stilbene structure is of the order of 3·5 A, similar to or slightly less than the values found for dibenzyl (compare Robertson 1935*a*). Only a few of these distances have been calculated for stilbene, but the shortest found is between *A* on molecule 1 and *E* on the reflected molecule 1 where the value is 3·58 A.

The centres of the four molecules in the unit cell of stilbene coincide with four crystal centres of symmetry at (000), ($\frac{1}{2}\frac{1}{2}0$), ($00\frac{1}{2}$) and ($\frac{1}{2}\frac{1}{2}\frac{1}{2}$), (Structure I). It was pointed out in the previous paper (Robertson, Prasad

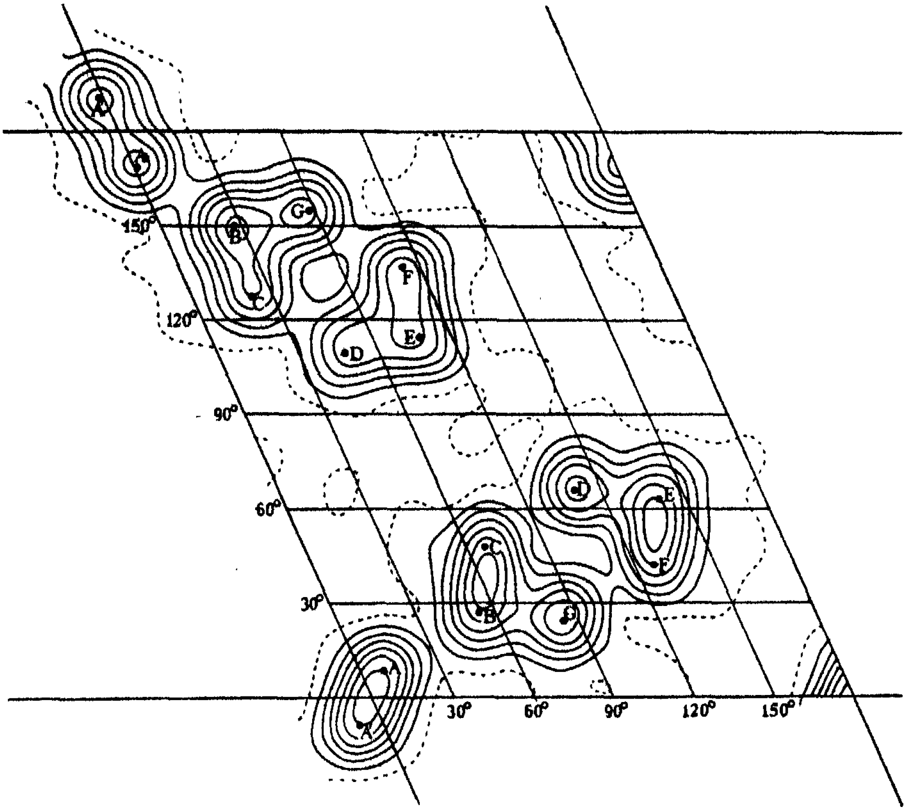


FIG. 6—The co-ordinates assigned to the atoms in stilbene.

TABLE V

hkl	*F* meas.	*F* calc. (I)	*F* calc. (II)
110	43	+51	−29
210	26	−29	+27
310	< 7	+ 1	+17
410	29	−27	+13
510	35	+33	− 3
610	<10	+ 5	+ 9
710	13	−12	0
810	<10	+15	+ 9
910	<10	+10	− 2
130	<10	−16	−20
230	23	+17	− 7
330	15	+ 9	+ 1
430	<10	+ 7	+17
530	<10	− 4	+20
630	<10	− 9	+ 3
730	<10	−10	−17
150	14	−20	+ 2
250	17	−16	− 4

and Woodward 1936) that a possible alternative arrangement would be (000), $(\frac{1}{2}\frac{1}{2}0)$, $(0\frac{1}{2}\frac{1}{2})$ and $(\frac{1}{2}0\frac{1}{2})$, (Structure II). In the second arrangement, the *b* axis projection and consequently the intensities of the (*h*0*l*) reflexions would be identical, the *a* axis projection and some of the (0*kl*) reflexions would be slightly changed, but the *c* axis projection and the (*hk*0) reflexions when *k* is odd would be considerably changed. The measured structure factors for these planes are compared with values calculated from the two alternative structures in Table V, and it will be seen that the first alternative is correct, the total discrepancies in the second case being nearly four times as great.

In conclusion, we wish to record our thanks to Sir William Bragg, O.M., P.R.S., for his continued interest in this work.

SUMMARY

The detailed structure of stilbene has been deduced from intensity measurements on the crystal by successive approximations involving double Fourier syntheses of the (*h*0*l*) zone of structure factors. The molecules are shown to be nearly planar and may be exactly planar within the limits of experimental error. This planar form is in striking contrast to the shape of the dibenzyl molecule in the crystal, where the benzene rings are almost at right angles to the plane of the central zig-zag. The planar form of the stilbene molecule is explained by the conjugation of the central double bond with the benzene rings, and the consequent contribution which an alternative excited structure makes to the normal state. The resonance involved is also shown by the thermochemical data and by the internuclear carbon-carbon distances obtained from the analysis, the "single bonds" emerging from the benzene rings having a length of only about 1·44 A. The length of the double bond is about 1·33 A, although this is not a very direct determination owing to difficulties in the analysis.

Two stilbene molecules contribute to the asymmetric unit of the crystal, and although these molecules have different appearances in the electron density maps the differences are shown to be due to the relative orientation of the molecules and not to any appreciable differences in their shape or dimensions. The precise orientation of each molecule is given, with the co-ordinates of the atoms. The minimum intermolecular approach distances are of the order of 3·5 A, similar to the distances found in other aromatic hydrocarbons.

REFERENCES

Pauling and Sherman 1933 *J. Chem. Phys.* **1**, 606.
Penney 1937 *Proc. Roy. Soc.* A, **158**, 306.
Robertson 1934 *Proc. Roy. Soc.* A, **146**, 473.
— 1935*a* *Proc. Roy. Soc.* A, **150**, 348.
— 1935*b* *Proc. Roy. Soc.* A, **150**, 106.
— 1936 *Nature, Lond.*, **138**, 683.
Robertson, Prasad and Woodward 1936 *Proc. Roy. Soc.* A, **154**, 187.

The Suspension of Solids in a Turbulent Stream

BY E. G. RICHARDSON, B.A., D.SC., *King's College, Newcastle-upon-Tyne*

(*Communicated by W. E. Curtis, F.R.S.—Received* 14 *July* 1937)

A considerable amount of research has been done recently on the mechanism by which grains of sand, etc., may be picked up by a river or by a gale of wind and carried to some distance. In particular Jeffreys (1929) has considered the theoretical aspects of the action while the writer (1934) and Bagnold (1936) have published experimental data on the action by which the grains are picked up in water and air respectively. Another problem of considerable practical importance is that of the quantity—or, more strictly, the time average of the quantity—of solid material carried at different levels above the bed or ground. Up to the present, data on this question have all been collected in relation to a special case, i.e. that in which the particles have the same size and density. Thus, Hurst (1929) collected samples of uniformly graded sand at different depths in a turbo-cylinder containing a vigorously stirred mixture of sieved sand in water, and showed that, in the body of the cylinder, an exponential relation between "weight" of sample and height above the bottom of the vessel existed.

The present paper is mainly concerned with measurements of the concentration of the suspended material near the uniformly graded bed of an artificial channel and the application of these results to the conditions in a natural stream.

THEORETICAL ASPECTS OF THE PROBLEM

A substance in finely divided form which would in still water fall to the bottom of the channel may remain in suspension when the water is in turbulent motion. The equilibrium for a level y where the silt density is maintained at a value s by the motion, has been formulated by Schmidt (1925) as follows, viz.

$$Cs + A\,ds/dy = 0, \tag{1}$$

where C is the velocity of free fall of the particles (given by Stokes's law or its recent modifications), and A is the *austausch* coefficient, and has the same dimensions as the diffusion coefficient. If A is independent of y, as it would be in a completely turbulent stream remote from a solid boundary, the solution of (1) is

$$A[\log s_0 - \log s] = Cy. \tag{2}$$

In the vicinity of the bed of the stream however, A will not remain constant, but will bear a close relation to the rate of diffusion of turbulence, as will appear in what follows. A is related to the friction drop by the equation of Prandtl (1933) and O'Brien (1933),

$$v_*^2 = \tau/\rho = A\,|dU/dy|, \tag{3}$$

where v_* has the dimensions of a velocity and is set equal to the geometric mean of the velocity fluctuations in the x and y directions, viz. to $\sqrt{(u'v')}$. Prandtl in a well-known hypothesis makes both of these fluctuations depend on the velocity gradient, so that in isotropic turbulence

$$u' = v' = l\,|dU/dy|, \tag{4}$$

where l is the factor of proportionality called the *mischungsweg* and U is the velocity of the stream at this point.

Kármán (1930) develops U as a Taylorian series to three terms. The only quantity having the dimensions of length which can be constructed from the second and third terms of his series involves $\frac{dU}{dy}\Big/\frac{d^2U}{dy^2}$, so he writes

$$l = x\left|\frac{dU}{dy}\Big/\frac{d^2U}{dy^2}\right|.$$

Various attempts have been made to investigate the variation of l across a wind channel by plotting the velocity distribution transverse to the axis, but the anemometers available are not really sensitive enough to enable one to measure the quotient of velocity gradient and curvature with sufficient precision for use in this formula. Beside the method of correlation of fluctuations of velocity developed by G. I. Taylor (1935–6) the present method of measurement of simultaneous silt concentration and velocity gradient enables one to calculate the variation of l across the stream. Thus from (2), the *austausch* coefficient

$$A = -Cs \Big/ \frac{ds}{dy},$$

but this, from (3) and (4), is also equal to

$$l^2 \left| \frac{dU}{dy} \right|. \tag{5}$$

Comparison of the velocity and concentration gradients at different depths in the stream will therefore enable one to determine the variation of l across a vertical section.

EXPERIMENTS IN AN ARTIFICIAL CHANNEL

The determination of velocity is made by a hot-wire anemometer. The concentration of suspended material is measured photoelectrically by the light absorbed from a limited beam crossing the channel horizontally at right angles to the walls. Some preliminary experiments to test the application of these instruments to silt problems have already been made by the writer (1934). These established the use of the method but were made in a miniature channel and at speeds scarcely attaining the critical value for turbulence so that they could not be applied to the problem now under discussion.

The new channel was constructed with plate glass sides 6 ft. long extending between two reservoirs of 10 cu. ft. capacity. The experimental portion was 1 ft. wide and allowed a maximum water depth of the same value. The two reservoirs were connected through a 3 in. pipe and a 2 h.p. centrifugal pump, so that the water after spilling over a weir in the downstream reservoir was circulated to the upstream one, where it again passed over an adjustable weir before re-entering the channel. Level gauges could be adjusted to touch the water surface at intervals.

For the mounting of the photoelectric cell and its associated beam of light it was essential that adjustment in both the horizontal (x) and vertical (y) directions should be possible without relative motion of the lamp and the cell. Fig. 1 shows a transverse section of the large channel with the apparatus for measuring silt concentration. Light from the 100 W lamp, *A*, passes through the circular and cylindrical lenses, *LL*, and the adjustable slit, S_1, and is so converted to a narrow beam which crosses the channel, passes through a second slit, S_2, and falls on the photoelectric cell. S_2 was adjustable for width of beam in the direction Ox; S_1 in both Ox and Oy. The vertical width of the beam was generally 2·0 mm.; the horizontal width 1 cm. By means of wheels *WW* rolling on rails, *RR*, the whole apparatus could be run along parallel to the axis of the channel. A handle rotating the long screw, *HH*, served to raise or lower the pillars, *PPPP*, through worm gears and so change the level of the whole apparatus relative to the channel without loss of relative location of the beam and the cell.

The apparatus was calibrated in the sense of establishing a relation between the concentration of the solid matter intercepting the beam of light and the defect of photoelectric current produced thereby. It was first established, by setting a Pointolite lamp at various distances from the cell (of the photronic type) and assuming the inverse square law, that the cell gave a linear response to the intensity of illumination over the range which it was contemplated to employ. It was then shown, by introducing a tank of water containing suspensions of solid particles of uniform size and at various concentrations into the path of the beam of light, that the extinction of light and hence the defect of photoelectric current was proportional (1) directly to the numbers of particles present, as between suspensions of particles of the same size but differing concentrations, (2) directly as the total projected area, i.e. to Σnd^2 when suspensions of particles of various but uniform diameters d were being compared. This law was verified for particles of diameters from 12μ to 2 mm. Before each set of silt readings, it was of course necessary to take a reading in clear water as "zero".

The hot wire for velocity measurements was of nickel, 0·002 in. diameter and 2 in. long. It was likewise calibrated by being mounted on a whirling arm working in a trough of water. It was mounted on a fork in the channel in such a way that it could be given vertical adjustment through a micrometer screw, while the whole anemometer could be moved along horizontally like the photoelectric apparatus. To save trouble in converting readings the voltage across the nickel wire due to the heating current was taken to the grid of a variable-μ valve, the corresponding current in the plate circuit being read on a sensitive galvanometer. With suitable grid bias it is then possible to secure a linear relationship between velocity and galvanometer reading. Both galvanometers employed in the measurements were damped in order that the readings obtained should represent a good average independent of casual fluctuations in concentration or velocity.

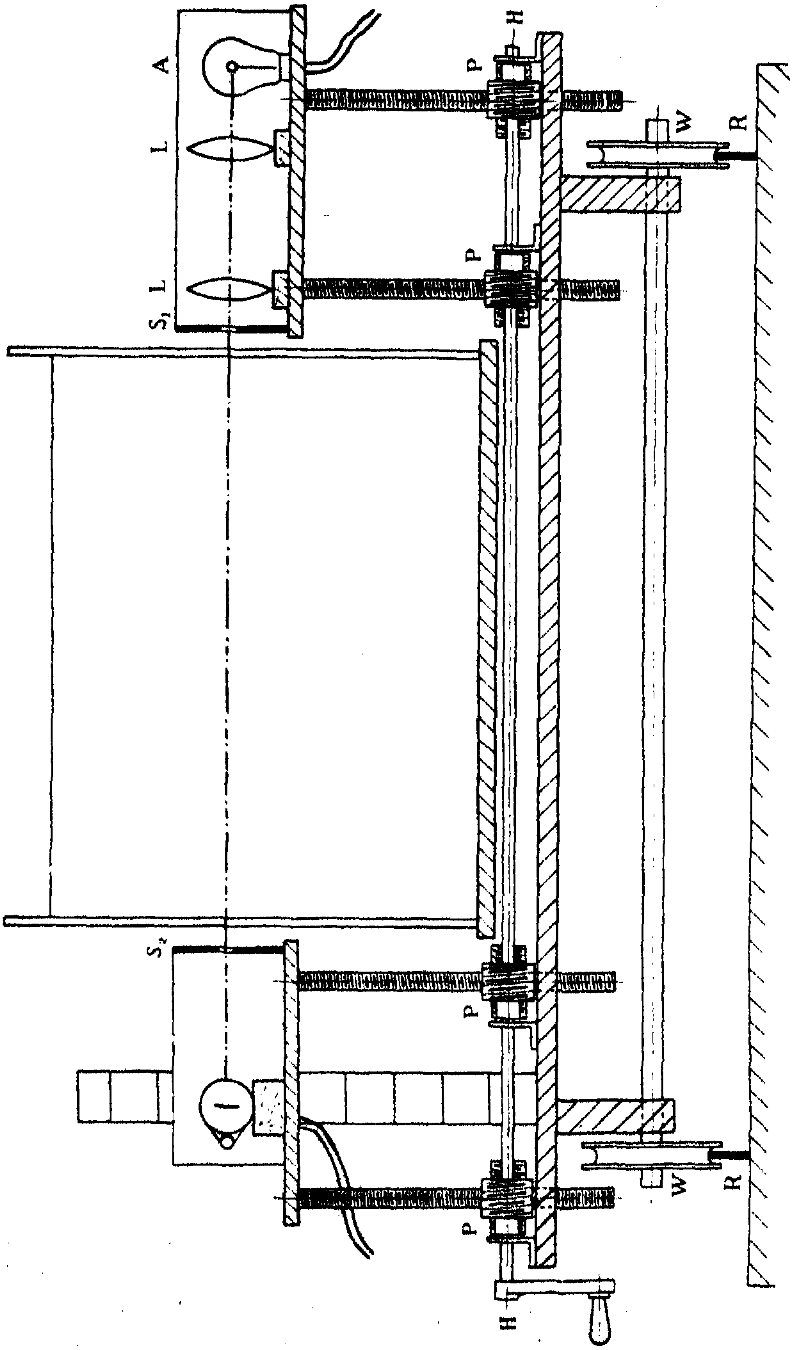


Fig. 1

Wind-blown sand was sieved into three sizes comprised within the following diameters: (1) 0·2–0·25 mm., (2) 0·25–0·3 mm., (3) 0·3–0·35 mm. By sedimentation further fractionation was carried out, and mechanical analysis of the resulting sands (for method *vide infra*) showed that the standard deviation from the mean diameter was about 0·01 mm. These pseudo mono-disperse sands had mean velocities of free fall in the channel water of 5, 7 and 10 cm./sec. respectively. Each specimen was evenly laid in turn to a depth of 3 mm. along the central portion of the floor of the channel for a length of 60 cm. The velocity gradient varied a good deal with x at the beginning of the artificial bed but both it and the sand distribution in the water were well established between 15 and 25 cm. from the head of the sand bed, and measurements of the vertical distribution of concentration and velocity were made over this region for two values of the mean channel speed. A set of measurements had to be completed within a few minutes as ripples would begin to form in time and spoil the readings. The method adopted was to get an assistant to raise the instruments through the water step by step while one noted the readings of the galvanometers. The need for such expedition set a limit to the damping which could be applied to the galvanometers, owing to the time taken for them to reach a steady reading. This ripple trouble was more blatant in the case of the finest sand. Results for these three sands at a mean channel speed of 45 cm./sec. are shown on figs. 2, 3 and 4. The velocity distributions (crosses) follow a logarithmic law down to a level a little above the top of the original bed where there is a point

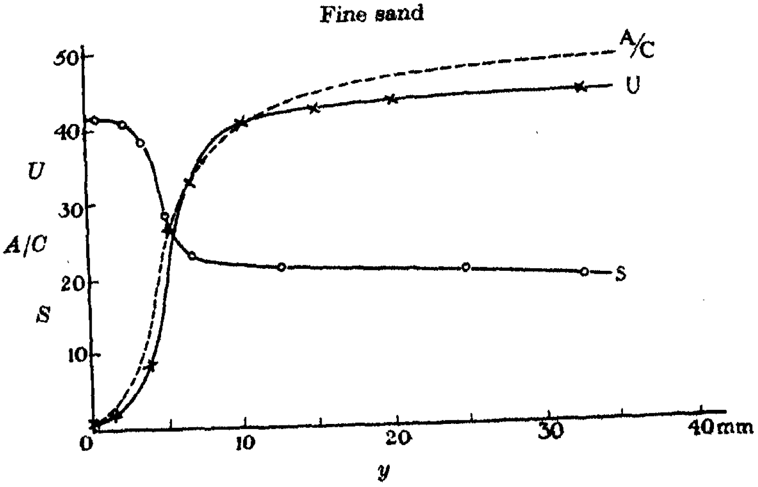


Fig. 2

of inversion. The suspension is very concentrated here and exhibits the anomalous viscosity common to colloids. It behaves like a semi-solid, a fact which is an impediment to the comparison of a hydrodynamic theory with experiment in this region.

Medium sand

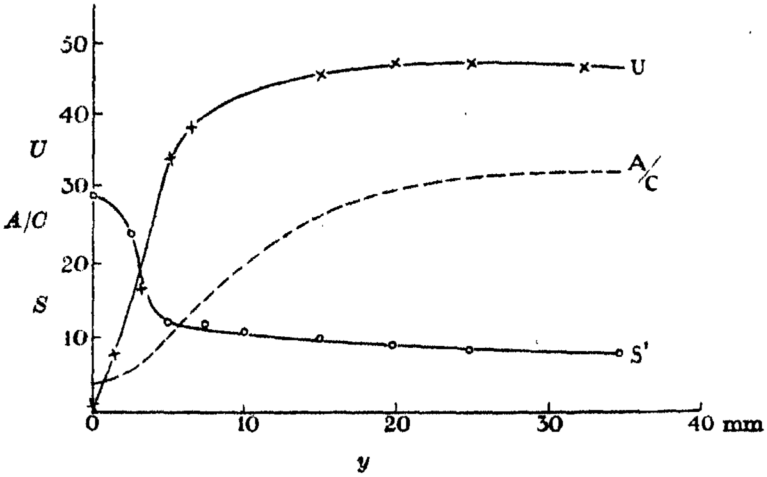


Fig. 3

Coarse sand

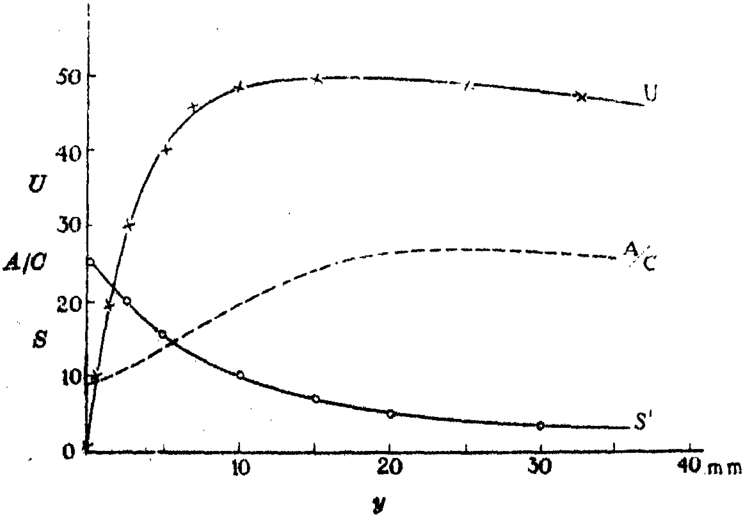


Fig. 4

The same three figures show as circles on continuous lines the solid concentration s and also that, as broken lines, of A/C from equation (2), i.e. of

$s(\partial s/\partial y)^{-1}$ with height above the channel floor. In each case after a small and nearly constant value through the mobile portion of the bed the *austausch* rises to reach a constant value a few centimetres above the bed. Multiplying by the appropriate value of C and dividing by the velocity gradient, we obtain the value of l^2 in accordance with equation (5). The values of the mixing length so obtained are plotted against the height above the channel floor on a logarithmic scale on fig. 5. It has been found impossible by logarithmic plotting or other means to cover the whole range of A or l by a power law. It seems best to group the results under three regions.

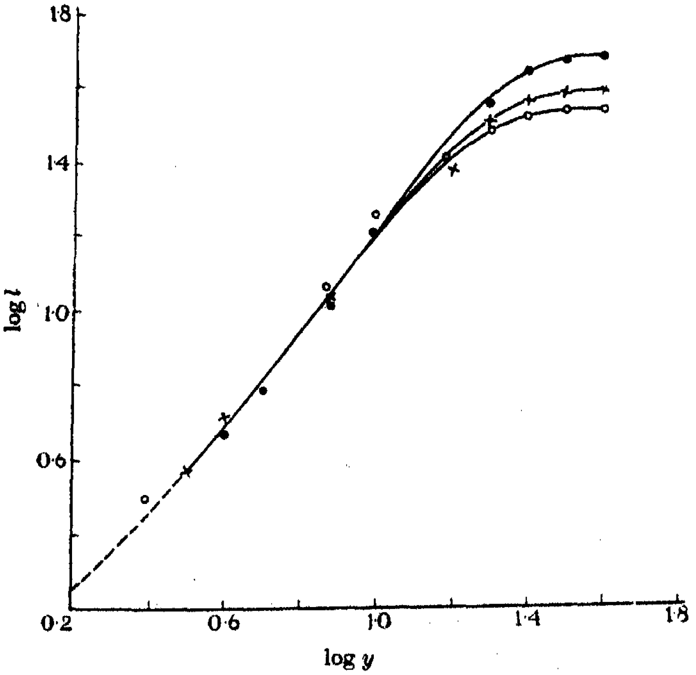


FIG. 5—● Coarse; × medium; ○ fine.

1—*Within the Bed*—Motion takes place to a limited extent under conditions approaching the laws of friction of solid bodies. The extent of the motion and its transmission depend on the permeability of the bed to water and the extent to which the particles are interlocked, i.e. on the compactness of the bed.

2—*In the Boundary Layer*—The system moves with increasing velocity and diffusion of the turbulence so engendered, following the frictional behaviour of colloidal and quasi-solid bodies in relative motion. Over this region, the following approximate relations between the factors hold:

$$U \propto \log(y+1);\ s \propto y^{-1};\ A \propto y;\ l \propto y.$$

The factor of proportionality between l and y will be a function of the relative "strength" of the turbulence in the stream and so, among other things, of the mean channel speed. In the present instance $l = 0{\cdot}3y$ approx. for all three grain sizes.

3—*In the Open Stream*—Velocity and mixing length are fully established and, if the absolute concentration of silt carried in suspension is not too great, the system flows like a true liquid. The relations are now

$$U, A \text{ and } l \text{ constant};\ |\log s| \propto y.$$

It will be noticed on fig. 5 that the steady value of l in the body of the stream is not the same for all three sands, being lower for the larger grains. This may indicate a diminution in the scale of the turbulence when the larger particles are carried, but it must be borne in mind that the value assigned to l in this region depends on that of C. It is well known to mining engineers in connexion with coal-washing processes that it does not do to assume that the velocity of fall of a particle in a maintained suspension of other particles will be the same as in unadulterated water. Actually, as mentioned earlier, we have determined the velocity of fall of a *cloud* of these particles in still water, but that does not resolve the doubt whether this is the correct value for the equilibrium of the system envisaged in equation (1), since we are not therein considering the equilibrium of particles in still water, but in turbulent streaming. A similar difficulty arises in predicting what size of particle will be held in suspension by an *upward* current of turbulent water in the elutriation process for separating minerals. Fortunately our imperfect knowledge of the value of C does not invalidate our deduction of the variation of A with depth, provided only that C can be considered constant for a given grain under the given degree of turbulence.

RIVER MEASUREMENTS

When we seek to apply the above theory and the results obtained in the artificial channel to natural streams, we must bear in mind that the solid matter carried in suspension by rivers differs from that which we have used up to the present in three respects: (1) it may consist of living as well as inorganic matter, (2) it is non-uniform in density, (3) it is non-uniform in size.

On the distribution of micro-plankton in the sea and in river estuaries much work has been done by Russell (1933).

It is well known that the intensity of the light from the sky which penetrates the water has a considerable effect on the vertical distribution of plankton, so that although this follows roughly the error law, the position of the maximum is liable to shift from day to day and from day to night. It is obvious then that any attempt to correlate the previous work with measurements in a natural stream cannot embrace the plankton and it is expedient to choose one in which the matter carried in suspension is mainly silt or detritus.

From this point of view, the River Tyne is admirable. Owing to pollution the estuary is markedly deficient in plankton (Jorgensen 1930) as compared with others which have not become so heavily industrialized. At the Swing Bridge, Newcastle, the river is restricted to a channel some 100 ft. wide and the bed is alluvium. Samples of the water with its silt content were taken at half-tide from one of the piers of the bridge where the deck overhung the water. The depth of the water is then about 20 ft., the tide is running out very fast and to visual appearance is very turbulent. There is also sufficient matter carried in suspension to render submerged objects invisible from above if they are more than a yard below the free surface.

The sampling was done by means of an apparatus kindly lent to the author by Dr. H. O. Bull of the Dove Marine Laboratory at Cullercoats. It consists of a lead-weighted copper container, the lid of which can be tightly sealed and has an inlet tube which runs down inside a specimen bottle within the container. The lid also has a vent to let out the air when the water enters by this tube. The vent was corked up until the container had been lowered to the desired depth, when the cork was removed by a string and the water with its solid matter entered the bottle with the minimum of commotion. This took about half a minute. Then the container was hauled up and the bottle removed. Samples were taken every 3 ft. down to the bed.

The velocity of the stream was measured at the same *stations* with an Ambler meter. This is a screw device, rotated by the water and ringing a bell every few revolutions, and had been previously calibrated in a water channel.

The river samples were first passed through a sieve to remove all objects larger than 2 mm. in diameter, and the remaining silt submitted to mechanical analysis in the author's own apparatus. This has been described in detail elsewhere (Richardson 1936), but the principles underlying the analysis may be recapitulated here. The idea of the apparatus is that after complete dispersion in water, the particles are allowed to settle in accordance with

Stokes's law and so become sorted out in such a fashion that a sample taken at a given depth after a given time will include in their original concentrations all those whose size is associated with a velocity insufficient to carry them beyond this depth within the time concerned. The sampling is not done literally but by measuring the amount of light absorbed from a narrow beam passing athwart the tank at the fixed level and falling on a photoelectric cell.

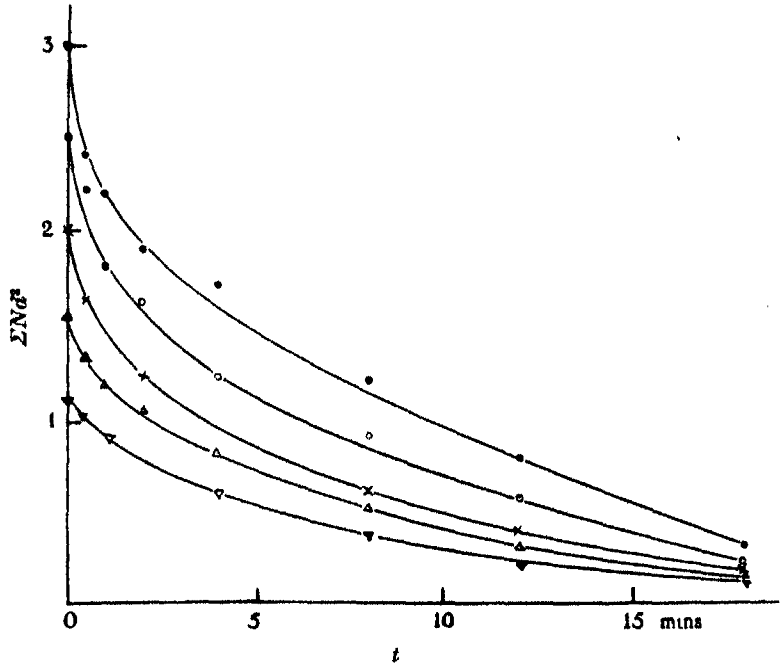


FIG. 6—▽ $y = 19$ ft.; △ $y = 14{\cdot}5$ ft.; × $y = 10$ ft.; ○ $y = 5{\cdot}5$ ft.; ● $y = 1$ ft.

For the coarser particles such as we are here concerned with, the absorption of light is proportional to the total surface they present, i.e. to

$$\sum_{d=0}^{d=\sqrt{y/kt}} Nd^2,$$

where N is the number of particles having a diameter d, k is the appropriate value of the factor relating the velocity of free fall with d^2 in Stokes's law, while y is the fixed depth and t the time. As the time goes on a curve of the light absorbed against the time can be constructed which is in fact the summation curve for the suspension. Such curves are given for the river samples on fig. 6. For given values of t (i.e. of d^2) the slopes of these curves will be proportional to $N_1 d_1^2$, where N_1 represents the number of particles for which

$d_1^2 = y/kt$. Such slopes are plotted in logarithmic form against depth in fig. 7 and justify by their straightness over the greater portion of the stream the assumption of nearly constant *austausch* (cf. equation (1)). Of course these measurements do not reach to the boundary layer above the bed.

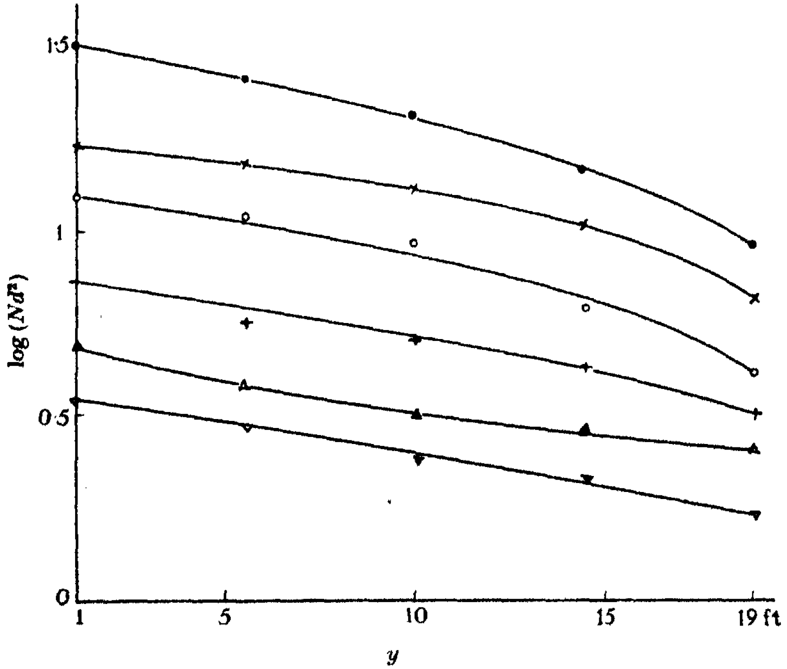


FIG. 7—● $d = 60\mu$; × $d = 42\mu$; ○ $d = 30\mu$; + $d = 20\mu$; △ $d = 15\mu$; ▽ $d = 10\mu$.

From the point of view of the river engineer it is the total quantity of material carried in suspension that matters, since this represents the destructive action of the river on its bed. The total silt content between $y = 0$ and $y = h$ (level of the free surface) can be estimated with sufficient approximation for practical purposes if the quantity carried by unit volume at two levels and the mean speed of the stream are known. Thus if we know the values σ_1 and σ_2 of $\log s$ at y_1 and y_2 respectively and postulate constant *austausch* over the mass of the stream, the total quantity of solid matter carried in a column of unit cross-section is given in logarithmic form by

$$S = \int_0^h \log s\,dy = \tfrac{1}{2}h(\sigma_0 + \sigma_h).$$

But
$$\sigma_0 = \sigma_1 + \frac{\sigma_1 - \sigma_2}{y_2 - y_1} y_1$$

and

$$\sigma_h = \sigma_1 - \frac{\sigma_1 - \sigma_2}{y_2 - y_1}(h - y_1),$$

whence

$$S = h\left[\sigma_1 + \frac{\sigma_1 - \sigma_2}{y_2 - y_1}(y_1 - \tfrac{1}{2}h)\right]. \tag{6}$$

This, multiplied by the mean speed of the stream, V, gives the amount transported in unit time across a rectangle perpendicular to the flow having unit base and height equal to the depth of the river. Hence the mean value of the erosion per unit area of the bed in the vicinity per unit time is SVR/P, where R is the hydraulic mean radius and P the wetted perimeter.

To facilitate the carrying out of routine experiments of this nature, the author has constructed a nephelometer of a type suitable for measuring relative silt concentrations in the stream itself. The instrument consists of a 12 V lamp at the focus of a lens in a watertight tube shining a beam on to a rectifier photoelectric cell, as used in the other experiments, at a distance of some 6 in. The space between is open to the water but the cell is shielded from the direct light of the sky. These two units are mounted in fixed positions on the cross-bar of a wooden T-piece, the other arm of which is loaded to keep the cross-bar horizontal in the water when the apparatus is submerged. The instrument was used in the Tyne on the same occasion that the samples were taken for analysis, in order to compare results. The observations of photoelectric defect I (on a portable microammeter) for various depths together with the velocity distribution are given on fig. 8. Except near the surface where scattered daylight was obviously penetrating the instrument the logarithmic distribution of the solid matter in suspension was confirmed. Two such instruments mounted at depths of about one- and two-thirds respectively of the total depth, with a velocity meter fixed half-way between, serve as a convenient permanent indicator of the hourly and diurnal silt content of the river, calculated in accordance with equation (6).

Finally, there is another problem of technical importance to which this work can lend evidence, i.e. that of reducing erosion whether it is caused by wind or by water. We have found that the absolute value of the *austausch* and therefore the extent of the erosion is a function of the scale of turbulence and so of the velocity gradient in the boundary layer. If this can be kept low without impeding the free passage of the main stream, our observations point to the planting of a sturdy vegetation which will grow to the height of a few inches above the soil as the obvious remedy. Such a crop will also resist erosion by consolidating the soil and making the individual grains less susceptible to removal.

The author wishes to thank Professors W. E. Curtis and G. I. Taylor for encouragement and criticism during the course of this work. He is also indebted to the Research Committee of Armstrong College for a grant to defray the cost of the channel and to Dr. C. M. White for allowing him to inspect the water channels at the City and Guilds College, London, before designing his own.

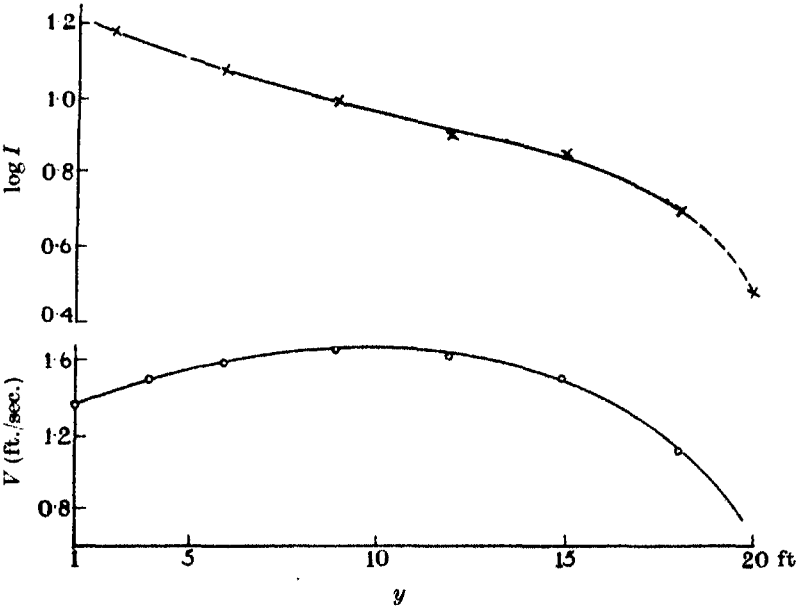


FIG. 8

SUMMARY

It is shown how the concept of an "*austausch* coefficient" in problems involving the diffusion of turbulence may be applied to the study of the distribution of solid material in a turbulent stream. A technique is developed by means of which the relative quantities of solid carried in suspension at different levels in a water channel having a bed of graded sand can be measured, and the results applied to predict how the *austausch* varies in the vicinity of the bed. The factor in question is found to increase in direct proportion to height at first but rapidly reaches a constant value in the main stream. Measurements in a natural river also indicate an approximately constant value, and, making this assumption of constant *austausch*, it is shown that a simple apparatus enables one to make observations of the total silt content of a river from time to time.

REFERENCES

Bagnold, R. A. 1936 *Proc. Roy. Soc.* A, **157**, 594–620.
Hurst, H. E. 1929 *Proc. Roy. Soc.* A, **124**, 196–201.
Jeffreys, H. 1929 *Proc. Camb. Phil. Soc.* **25**, 272–6.
Jorgensen, O. M. 1930 *Proc. Univ. Durham Phil. Soc.* **8**, 41–54.
Kármán, Th. von 1930 *Nachr. Ges. Wiss. Göttingen*, pp. 58–76.
O'Brien, M. P. 1933 *Trans. Amer. Geophys. Soc.* pp. 487–91.
Prandtl, L. 1933 *Z. Ver. dtsch. Ing.* **77**, 105–14.
Richardson, E. G. 1934 *Phil. Mag.* **17**, 769–83.
—— 1936 *J. Sci. Instrum.* **13**, 229–33.
Russell, F. S. 1933 *J. Mar. Biol. Ass. U.K.* **19**, 569–81.
Schmidt, W. 1925 "Die Massenaustausch in freier Luft." Hamburg.
Taylor, G. I. 1935 *Proc. Roy. Soc.* A, **151**, 444–54.
—— 1936 *Proc. Roy. Soc.* A, **157**, 537–46.

INDEX TO VOLUME CLXII (A)

Absorption spectra of monosulphides (Mathur), 83.
Absorption spectrum of formic acid (Price and Evans), 110.
Acetylene, slow combustion (Bone and Carruthers), 502.
Amine series, rotatory dispersion in (Lowry and Baldwin, and Baldwin), 204, 215, 228.
Appleton (E. V.) The Bakerian Lecture—Regularities and irregularities in the ionosphere—I, 451.
Atmospherics, nature of—V (Watt, Herd and Lutkin), 267.

Bakerian Lecture (Appleton), 451.
Baldwin (W. C. G.) Rotatory dispersion in the amine series. II—*l*-secondary butylamine and its derivatives, 215.
Baldwin (W. C. G.) Rotatory dispersion in the amine series. III—The absorption spectra of diamines, 228.
Baldwin (W. C. G.) *See* Lowry and Baldwin.
Bard (B. J. A.) *See* Bone and Bard.
Boardman (L. E.) and Wild (N. E.) The diffusion of pairs of gases with molecules of equal mass, 511.
Bone (W. A.) and Bard (B. J. A.) Researches on the chemistry of coal. VIII—Comparative studies of the macro-constituents of bituminous coals, 495.
Bone (W. A.) and Carruthers (J. E.) The slow combustion of acetylene, 502.
Bosworth (R. C. L.) and Rideal (E. K.) Studies in contact potentials. The condensation of potassium and sodium on tungsten, 1.
Bosworth (R. C. L.) Studies in contact potentials. The evaporation of sodium films, 32.
Briggs (G. H.) *See* Goldhaber and Briggs.
Bromine, liquid, absorption spectrum of (Porret), 414.

Carruthers (J. E.) *See* Bone and Carruthers.
Chalmers (B.) The influence of the difference of orientation of two crystals on the mechanical effect of their boundary, 120.
Chlorophyll, reversible oxidation of (Rabinowitch and Weiss), 251.
Coal, chemistry researches (Bone and Bard), 495.
Collens (H.) *See* Malan and Collens.
Contact potentials, studies in (Bosworth and Rideal, and Bosworth), 1, 32.
Conway (A. W.) Quaternion treatment of the relativistic wave equation, 145.
Cooke (A. H.) and Hull (R. A.) Experiments on iron ammonium alum in alternating fields at very low temperatures, 404.
Crystal boundary, mechanical effect (Chalmers), 120.

Dibenzyl series, X-ray analysis of (Robertson and Woodward), 568.
Diethyl ether, influence of hydrogen and deuterium on the thermal decomposition of (Mitchell and Hinshelwood), 357.
Diffusion of pairs of gases (Boardman and Wild), 511.

Eddington (Sir Arthur) Theory of scattering of protons by protons, 155.
Eggleston (F. F. H.) and Martin (L. H.) The angular distribution of photo electrons from the *K* shell, 95.
Elastic stability of a rectangular plate (Leggett), 62.
Electronic specific heat and X-ray absorption of metals (Jones and Mott), 49.
Ethane, inhibition by nitric oxide of the decomposition of (Staveley), 557.
Evans (W. M.) *See* Price and Evans.

Foster (J. S.) and Snell (H.) The stark effect in hydrogen and deuterium, 349.
Foster (J. S.) *See* Panter and Foster.
Fox (J. J.) and Martin (A. E.) Investigations of infra-red spectra—absorption of some hydroxy compounds in the region of 3μ, 419.

Goldhaber (M.) and Briggs (G. H.) Scattering of slow neutrons, 127.

Herd (the late J. F.) *See* Watt, Herd and Lutkin.
Hinshelwood (C. N.) *See* Mitchell and Hinshelwood.
Hull (R. A.) *See* Cooke and Hull.

Infra-red spectra, investigations of (Fox and Martin), 419.
Iron ammonium alum in alternating fields, experiments on, at very low temperatures (Cooke and Hull), 404.

Jeffreys (H.) The tests for sampling differences and contingency, 479.
Jones (H.) and Mott (N. F.) The electronic specific heat and X-ray absorption of metals, and some other properties related to electron bands, 49.

Kothari (D. S.) Neutrons, degeneracy and white dwarfs, 521.

Leggett (D. M. A.) The elastic stability of a long and slightly bent rectangular plate under uniform shear, 62.
Leigh-Smith (Alice) *See* Richardson and Leigh-Smith.
Lowry (the late T. M.) and Baldwin (W. C. G.) Rotatory dispersion in the amine series. I—*d*-*Spiro*heptane-Diamine and its derivatives, 204.
Lutkin (F. E.) *See* Watt, Herd and Lutkin.

Makinson (R. E. B.) Metallic reflexion and the surface photoelectric effect, 367.
Malan (D. J.) and Collens (H.) Progressive lightning. III—The fine structure of return lightning strokes, 175.
Martin (A. E.) *See* Fox and Martin.
Martin (L. H.) *See* Eggleston and Martin.
Massey (H. S. W.) and Mohr (C. B. O.) Nuclear excitation and disintegration collisions involving strong interaction—I, 527.
Mathur (L. S.) The absorption spectra of the monosulphides of alkaline earth elements and their latent heats of vaporization, 83.
Meksyn (D.) Solution of Oseen's equations for an inclined elliptic cylinder in a viscous fluid, 232.
Michiels (J. L.) and Saxton (J. A.) The influence of temperature on the capture of neutrons by various elements, 441.

Mitchell (J. W.) and Hinshelwood (C. N.) The influence of hydrogen and deuterium on the thermal decomposition of diethyl ether in the low pressure region, 357.
Mohr (C. B. O.) *See* Massey and Mohr.
Mott (N. F.) *See* Jones and Mott.

Neutrons, degeneracy and white dwarfs (Kothari), 521.
Neutrons, influence of temperature on the capture of (Michiels and Saxton), 441.
Nuclear excitation (Massey and Mohr), 529.

Oseen's equations for an inclined elliptic cylinder, solution of (Meksyn), 232.
Osglim lamp, behaviour of—I, II (Richardson), 293, 316.

Panter (S. F.) and Foster (J. S.) Stark effect in iron and the contrast with pole effect, 336.
Photoelectric effect and metallic reflexion (Makinson), 367.
Photoelectrons, angular distribution of, from the K shell (Eggleston and Martin), 95.
Porret (D.) The absorption spectrum of liquid bromine, 414.
Potassium and sodium, condensation of, on tungsten (Bosworth and Rideal), 1.
Price (W. C.) and Evans (W. M.) The absorption spectrum of formic acid in the vacuum ultra-violet, 110.
Progressive lightning—III (Malan and Collens), 175.

Quaternion treatment of the relativistic wave (Conway), 145.

Rabinowitch (E.) and Weiss (J.) Reversible oxidation of chlorophyll, 251.
Richardson (E. G.) The suspension of solids in a turbulent stream, 583.
Richardson (H. O. W.) and Leigh-Smith (Alice) The shape of the continuous β-spectrum of thorium C.C′, 391.
Richardson (L. F.) The behaviour of an osglim lamp. I—Introduction, drift, equilibrium and A.C. bridge measurements. II—Oscillograph methods, steady motion and relaxational oscillations, 293, 316.
Rideal (E. K.) *See* Bosworth and Rideal.
Robertson (J. M.) and Woodward (I.) X-ray analysis of the dibenzyl series. IV—Detailed structure of stilbene, 568.

Sampling tests for differences and contingency (Jeffreys), 479.
Saxton (J. A.) *See* Michiels and Saxton.
Scattering of protons by protons (Eddington), 155.
Slow neutrons, scattering (Goldhaber and Briggs), 127.
Snell (H.) *See* Foster and Snell.
Sodium films, evaporation of (Bosworth), 32.
Stark effect in hydrogen and deuterium (Foster and Snell), 349.
Stark effect in iron (Panter and Foster), 336.
Staveley (L. A. K.) The inhibition by nitric oxide of the decomposition of ethane, 557.
Stilbene, detailed structure of (Robertson and Woodward), 568.
Suspension of solids (Richardson), 583.

Thorium C.C′, shape of the continuous β-spectrum of (Richardson and Leigh-Smith), 391.

Watt (R. A. W.), Herd (the late J. F.) and Lutkin (F. E.) On the nature of atmospherics—V, 267.
Weiss (J.) *See* Rabinowitch and Weiss.
Wild (N. E.) *See* Boardman and Wild.
Woodward (I.) *See* Robertson and Woodward.

END OF THE ONE HUNDRED AND SIXTY-SECOND VOLUME